

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, us-10-780-703-2.p2n.rge

Run on: April 6, 2006, 09:51:12 ; Search time 6373 Seconds

(without alignments)
3380.459 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAIVSGTTSIIIPMSQVNNNN.....MNPPIVDPGNDRHYTNIPFK 379

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/abse/ABSWEB_spool/US10780703/runat_06042006_101337_6322/app_query.fasta.1
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -ALIGN=15 -MODE=LOCAL
-USER=US10780703 @CN 1.1 4939 @runat_06042006_101337_6322 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_scs.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2034	100.0	1171 15	BT000874 Arabidops
2	1735	85.3	74903 15	AC005312 Arabidops
3	1574	77.4	1062 15	AF360308 Arabidops

4	815.5	40.1	1631	15	AK102902	AK102902 Oryza sat
5	796.5	39.2	1640	15	AK072275	AK072275 Oryza sat
6	796.5	39.2	1891	15	AK071052	AK071052 Oryza sat
7	789.5	38.8	2424	15	AK068393	AK068393 Oryza sat
8	781	38.5	1406	15	AY672069	AY672069 Hordeum v
9	781	38.4	1212	6	AX654050	AX654050 Sequence
10	769.5	37.8	80577	14	AP007649	AP007649 Lotus cor
11	769.5	37.8	91718	14	AP007409	AP007409 Lotus cor
12	759	37.3	1250	6	AX654570	AX654570 Sequence
13	688.5	33.8	549	6	AX653764	AX653764 Sequence
14	672.5	33.1	110000	15	AP008207_385	Continuation (386
15	672.5	33.1	110000	15	AP008207_386	Continuation (387
16	672.5	33.1	150064	15	AP003346	AP003346 Oryza sat
17	672.5	33.1	191022	15	AP003431	AP003431 Oryza sat
18	655.5	32.2	110000	15	AP008211_203	Continuation (204
19	655.5	32.2	173074	15	AC137623	AC137623 Oryza sat
20	611	30.0	110000	15	AP008214_007	Continuation (8 of
21	611	30.0	185545	15	AP005657	AP005657 Oryza sat
22	550	27.0	1271	15	AK107090	AK107090 Oryza sat
23	546.5	26.9	1591	15	AK073013	AK073013 Oryza sat
24	544.5	26.8	1158	6	AX654270	AX654270 Sequence
25	536	26.4	950	6	AX653803	AX653803 Sequence
26	517.5	25.4	1152	6	AX652795	AX652795 Sequence
27	517	25.4	1423	15	AY573802	AY573802 Lycopersi
28	516	25.4	1152	15	DQ028772	DQ028772 Glycine m
29	515	25.3	978	15	AF361804	AF361804 Arabidops
30	514	25.3	830	6	AX651430	AX651430 Sequence
31	514	25.3	1021	15	DQ028774	DQ028774 Glycine m
32	512	25.2	1567	15	AF011555	AF011555 Lycopersi
33	512	25.2	1664	15	AJ704829	AJ704829 Prunus pe
34	509	25.0	1209	6	AR569436	AR569436 Sequence
35	509	25.0	1209	6	AX087867	AX087867 Sequence
36	509	25.0	1565	15	AY704177	AY704177 Solanum t
37	508.5	25.0	972	15	BT008346	BT008346 Arabidops
38	508.5	25.0	1125	6	CS137954	CS137954 Sequence
39	508.5	25.0	1336	15	AK117878	AK117878 Arabidops
40	508	25.0	894	6	CS137686	CS137686 Sequence
41	508	25.0	925	15	AY091428	AY091428 Arabidops
42	508	25.0	1150	15	BT000790	BT000790 Arabidops
43	508	25.0	1230	15	AF428375	AF428375 Arabidops
44	508	25.0	1236	15	AF083738	AF083738 Arabidops
45	508	25.0	1237	15	AY087772	AY087772 Arabidops

ALIGNMENTS

RESULT 1	BT000874	Arabidopsis thaliana clone C104984 putative NAM (no apical meristem) protein (Atg902450) mRNA, complete cds.
LOCUS	BT000874	1171 bp mRNA linear PLN 16-OCT-2002
DEFINITION	Arabidopsis thaliana (thale cress)	
ACCESSION	BT000874.1	GI:24030185
VERSION	BT000874.1	GI:24030185
KEYWORDS	FLI CDNA.	
SOURCE	Arabidopsis thaliana	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi	
REFERENCE	1 (bases 1 to 1171)	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
AUTHORS	Arabidopsis Open Reading Frame (ORF) Clones	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 1171)	Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT Annotation based on January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES
Location/Qualifiers
1..1171
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="C104984"
/ecotype="Columbia"
/note="this clone is in pUNI 51."

gene
1..1171
/gene="At2g02450"

CDS
1..1140
/gene="At2g02450"
/codon_start=1
/evidence="experimental"
/product="putative NAM (no apical meristem) protein"
/protein_id="AA041274.1"
/db_xref="GI:24030186"
/translation="MAIVSSTTSIIIPMSNOVNNKEGIEDNHRGGQESHVONEDAD
DHDHVMGPRFPHTEBELIEFYLRKVEGKRFNVELITFLDLRYDPWELPMAAI
GEKWFYVPRDRKVRNGDRPNRTVSGYWKATGADMRISRTSPIGLKLTVYSG
KAPKRTSWINNEIRLPHETEKYKAEISLCRVYKRPVEDHPSVPSRSLRHHNH
NSSTSRALRQQHSSSHSDNNNNNNNNINLEKLSYEYSGDSTTTTTTNSNS
DVTIALANONIYRPMPTVSNNTLIVSTRNHQDDDEITAVDQLRVNYQISDGAATL
MPOTQALAMAMIPAGTIPNNALWDMNPVDPGNRDRHTNIPPK"
1141..1171
/gene="At2g02450"

3'UTR
misc_difference 1150
/gene="At2g02450"
/note="compared to genomic sequence"
/replace="a"

ORIGIN
Alignment Scores:
Pred. No.: 6,85e-150 Length: 1171
Score: 2034.00 Matches: 379
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-780-703-2 (1-379) x BT000874 (1-1171)

QY 1 MetAlaIleValSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
DB 1 ATGGCAATTGTATCTCCACCAAGCATCATTCCTCATGATGTAACCAAGTCAACATAAC 60
QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
DB 61 GAABAAGGTATAGACACATCATCATAGAGGGCGGCAAGAGAGTATCTCCAAATGAA 120
QY 41 AspGluAlaAspAspHisAspMetValMetProGlyPheArgPheHisProThr 60
DB 121 GATGAAGCTGATGATCATCATCATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 61 GluGluGluLysGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
DB 181 GAAGAAGAACTCATAGAGTTTACCTTCGCCGAAAGATTGAAGGCAACAGCTTTAATGTA 240
QY 81 GluLeuIleThrPheLeuAspLeuThrArgThrAspProThrGluLeuProAlaMetAla 100
DB 241 GAACATCATCATCTTCTCTCATCTTATCTGCTATGATCTTGGGAACTTCTGCTATGGCG 300
QY 101 AlalGlyGluLysGluThrPheThrValProArgAspArgLysThrArgAsnGly 120
DB 301 GCGATAGGAGAGAAAGAGTGGTACTTCTATGTCGCAAGAGATCGGAAATATAGAAATGA 360
QY 121 AspArgProAsnArgValThrThrSerGlyThrThrPheAlaThrGlyAlaAspArgMet 140

DB 361 GATAGACCAACCGAGTAAACGACTTCAGATATTGGAAAGCCACCGAGCTGATAGGATG 420
QY 141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
DB 421 ATCAGATCGAGAGACTTCTCGGCTATCGGCTATCGGCTATCGGCTATCGGCTATCGGCT 480
QY 161 LysAlaProLysGlyThrArgThrSerThrIleMetAsnGluTyrArgLeuProHis 180
DB 481 AAAGCCCTTAAAGGCACCTCGTACTAGTTGGATCATGAACGAGTATCGTCTTCGACCAT 540
QY 181 GluThrGluLysThrGlnLysAlaGluLeuSerLeuCysArgValTyrLysArgProGly 200
DB 541 GAACCGAGAGAGTACCAAAAGGCTGAAATATCATTTGTCCGAGTGTACAAAAGGCCAGCA 600
QY 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisAsnHisAsnSer 220
DB 601 GTAGAAGATCATCCATCGTACCAAGCTTCTCTCCACAAAGCATCATAAACATCACTCA 660
QY 221 SerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSer 240
DB 661 TCGACATCATCCGTTTACGCTTAAACAACAACCAACCAACCAACCAACCAACCAACCA 720
QY 241 SerAspAsnAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 260
DB 721 TCCGACACCAACCTTAAACAACAACAACAACAACAACAACAACAACAACAACAACA 780
QY 261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 280
DB 781 TATTCGGCGAGCGGAGCAGCAACAACAACAACAACAACAACAACAACAACAACAACA 840
QY 281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu 300
DB 841 GCTCTAGCCCAATCAAAACATATATCGTCCATGCTTACGACCAAGCAACAACAACA 360
QY 301 IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGln 320
DB 901 ATAGTCTCTACGAGAAATCATCAAGACGATGATGAACTGCCATTTGTTGACGATCTTCA 960
QY 321 ArgLeuValAsnTyrGlnIleSerAspGlyAlaThrThrLeuMetProGlnThrGlnAla 340
DB 961 AGACTAGTTAACTCAAAATATCATGATGAGCGAGCAACGCTAATGCTCAAACTCAAGCG 1020
QY 341 AlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuThrAspMet 360
DB 1021 GCGTTAGCTATGAACATGATCTCTCGAGAAAGATTCGAACATGCTTTGGGATG 1080
QY 361 TrpAsnProIleValProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
DB 1081 TGGATCCCAATAGTACCAAGATGGAACAGAGATCACTATATATCTCTTTTAAG 1137

RESULT 2
AC005312 74903 bp DNA linear PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone T16F16 map C1C11A04,
DEFINITION complete sequence.
AC005312
VERSION AC005312.3 GI:20197334
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 74903)
AUTHORS Rounsley,S.D., Lin,X., Kaul,S., Kaul,T.P., Fujii,C.Y., Mason,T.M.,
Shen,M., Renning,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 74903)
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

REFERENCE	3 (bases 1 to 74903)		lectin"	
TITLE	Town.C.D. and Kaul.S.		/pseudo	
AUTHORS	Direct Submission		/complement(10262..11417)	
JOURNAL	Submitted (27-FEB-2002) The Institute for Genomic Research, 9712	gene	/gene="At2g02250"	
COMMENT	Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org	mRNA	/note="synonym: T16F16.4"	
FEATURES	On Apr 18, 2002 this sequence version replaced gi:6598458.		/complement(join(<10262..10691,10775..10896,11052..11417))	
source	Location/Qualifiers	CDS	/gene="At2g02250"	
	1..74903		/complement(join(10262..10691,10775..10896,11052..11417))	
	/organism="Arabidopsis thaliana"		/gene="At2g02250"	
	/mol_type="genomic DNA"		/codon_start=1	
	/db_xref="taxon:3702"		/product="lectin-like protein"	
	/chromosome="2"		/protein_id="AAC78509.1"	
	/map="C1C11A04"		/db_xref="GI:3894159"	
	/clone="T16F16"		/translation="MGHKQVDSRGKRGKVPSSSMVQKHVSSGGAIISGSPSLFDN	
	/ecotype="Columbia"		/translating="SPRACVAASVSKTFESAVNSDSVMDKFLPSDYSLVPSRVSF	
gene	complement(<1..2731)		/translating="SPRACVAASVSKTFESAVNSDSVMDKFLPSDYSLVPSRVSF	
	/gene="At2g02220"		/translating="SPRACVAASVSKTFESAVNSDSVMDKFLPSDYSLVPSRVSF	
mRNA	/note="synonym: T16F16.1"		/translating="SPRACVAASVSKTFESAVNSDSVMDKFLPSDYSLVPSRVSF	
	complement(<1..2731)		/translating="SPRACVAASVSKTFESAVNSDSVMDKFLPSDYSLVPSRVSF	
	/gene="At2g02220"		/translating="SPRACVAASVSKTFESAVNSDSVMDKFLPSDYSLVPSRVSF	
CDS	complement(<1..2731)	gene	/complement(12061..13367)	
	/gene="At2g02220"		/gene="At2g02260"	
	/codon_start=1		/note="synonym: T16F16.5; pseudogene, reverse	
	/product="putative protein kinase"		transcriptase"	
	/protein_id="AAC78507.3"		/pseudo	
	/db_xref="GI:2019735"	gene	/complement(13604..15132)	
	/translation="MRVHRCVIVIFLTELCPYSSSQTTSRCHPHDLNLRDFA		/gene="At2g02270"	
	HLEPDPGWNSSSDTCGWTGTCNSNNTGRVIRLEGNKLSGKLSGLKDEI		/note="synonym: T16F16.6"	
	RVLNRFNFTIDPSIFNKNLQTLDSNDLSSGIGTPTINPALQSFDSLNKFN	mRNA	/complement(join(<13604..13663,13728..14100,14227..14312,	
	GSLSHCHNSTQIRVVKLAVNYFAGNFTSGFKCVLLEHLCLMDLTGNIPIRDLF		14854..>15132))	
	LKRNLLGQIKNSLGSIRLIRLNSLVLRLDWNLFSGEIPDVFDELQKFLPG	CDS	/gene="At2g02270"	
	TRFTGIGI PKSLANPSLNLNLNLSLGRMLNCTAMLNLSLDTGNFRNGRLPE		complement(join(13604..13663,13728..14100,14227..14312,	
	NLPDCKLKNVNLARNTHGQVPSFKNFESLSYFSLNSLANISSALGTLQCKNL		14854..15132))	
	LTVLTLNFGHLPDSSLFHFKLVVAVNCLTGMPSLSSNELQLDLUSWR		/gene="At2g02270"	
	LTGAPWIGDKALFYDLDSNFTGEIPKSLTKLESLSRNLISVNEPSPDFPMK		/codon_start=1	
	RNSARALQYNQIFGFTPIELGNHLSGPTWEEFKNLKLHVFDLKWNALSGIPSS		/product="putative phloem-specific lectin"	
	LSGNTLSLQNLNLSLGSIPVSLQSLFSLKFSVAYNLISVPSGQFQTPPNS		/protein_id="AAC78510.1"	
	FENHLCGEHFPCEGESALIKRSRSGDGMGIAFGSVFLLTLNLVLA		/db_xref="GI:3894160"	
	RRSCEVDYKATEESNMRKELGKLVVLFGSNDELSDYDLNLSNFDQANIL		/translation="MAQGEIVPRSPDGLPENCISNIISFTPRDACFAASVSKAFE	
	CGGFGVYKATPDGKKVAIKLUSGCGQIFERFVETLSRAQHLVLLKGFCE		/translating="MAQGEIVPRSPDGLPENCISNIISFTPRDACFAASVSKAFE	
	YKNDRLIYSYMGNSLDYMLHNRNDGPALKWTRLRIAQGAAGLLYLHGGCDPHI		/translating="MAQGEIVPRSPDGLPENCISNIISFTPRDACFAASVSKAFE	
	LHRDLSKNTLLDNFNLSHLADFLGLRLMSPYETHVSTDLVGTILGYPPEYGAQ		/translating="MAQGEIVPRSPDGLPENCISNIISFTPRDACFAASVSKAFE	
misc_feature	complement(1..1865)	gene	/complement(16098..17046)	
	/gene="At2g02220"		/gene="At2g02280"	
	/note="overlap with BAC clone F504 (AC005936:1..1865)."		/note="synonym: T16F16.7"	
repeat_region	3226..3247	mRNA	/complement(join(<16098..16334,16849..>17046))	
	/rpt_family="AT rich"		/gene="At2g02280"	
repeat_region	complement(3504..3524)		/complement(join(16098..16334,16849..17046))	
	/rpt_family="AT rich"	CDS	/gene="At2g02280"	
repeat_region	complement(4101..4133)		/codon_start=1	
	/rpt_family="AT rich"		/product="putative phloem-specific lectin"	
gene	complement(6000..7202)		/protein_id="AAC78511.1"	
	/gene="At2g02230"		/db_xref="GI:3894161"	
	/note="synonym: T16F16.2"		/translation="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
mRNA	complement(join(<6000..6237,6295..6555,6649..6767,		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	6867..7202))		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/gene="At2g02230"		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
CDS	complement(join(6000..6237,6295..6555,6649..6767,		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	6867..7202))		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/gene="At2g02230"		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/codon_start=1		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/product="putative phloem-specific lectin"		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/protein_id="AAC78508.1"		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/db_xref="GI:3894158"		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/translation="MEQIHGDSNGSGGGGGGSRNDEISVTRASRFDPEDICISKVI		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	SHTSPRACVAVSVKSAQSDLVWEMFLPSEYSLVLOSANHLKKEIFLSLAD		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	NSVLVNGKSGFWKEKSGKCYMLSAMELTIJWGDSPAYWKWITVPESKVEALR		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	NVCFEVRKISCSGLKGYHYVYVFKTANGSYGFDLPVPEAGVGVGVKVAIKS		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	VYFSGNADRSATSHSYGISEEVEGERGEMVNVGPKRVDGWSVEVLGKGFYIN		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	NGCGDDGSDSIEISIMETQGNWKSGLIIQIGIEIRPERSN"		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	complement(7966..9209)		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
gene	/gene="At2g02240"		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	
	/note="synonym: T16F16.3; pseudogene, phloem-specific		/translating="MNTQISQKTRYSAYIVYKTYRPHGFKHGVGFIHGTTPKAKR	

[illegible]


```
/product="putative NAM protein"  
/protein_id="AAK26018.2"  
/db_xref="GI:21450867"  
/translation="LYRDPWELPMAAIGKEWYFVPRDKYRNGRDRPNRVTTSGY  
WKATGADRMISERSPRLGLKTLVFGSKAPKGRTRTSMINVEYRLPHETEKYQKAE  
ISLCVYRPGVEDSPVSRSLSTRHHNSSTSSRLALRQOHHSSNSNNDNLNN  
NNNINLEKLSLEYSGDSTTTTNTNSNDVTIALANQNIYRPMPTDTSNNTLIVSTR  
NHODDETAIVDDLRQVYQISDGAATLMPQTAALAMNMPAGTI PNNALWDMNP  
IVPDGNRDHYTNIPK"  
882..1062  
/gene="At2g02450"
```

3'UTR

ORIGIN

Alignment Scores:

```
Pred. No.: 6 12e-114 Length: 1062  
Score: 1574.00 Matches: 292  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 77.4% Indels: 0  
DB: 15 Gaps: 0
```

US-10-780-703-2 (1-379) x AF360308 (1-1062)

```
Qy 88 LeuTyArgTyArgProTyrLeuProAlaMetAlaAlaLeuGlyGluLysGluTyr 107  
Db 3 CTTTATCGTATGATCTCTGGCACTCTCTGCTATGCGCGGATAGGAGAGAGAGTGG 62  
Qy 108 TyrPheTyValProArgAspArgLysTyArgAsnGlyAspArgProAsnArgValThr 127  
Db 63 TACTTCTATGTGCCAAGAGATCGGAATATAGAAATAGAGATAGACCGAACCGAGTAAAG 122  
Qy 128 ThrSerGlyTyTrpLysAlaThrGlyAlaAspArgMetLeuArgSerGluThrSerArg 147  
Db 123 ACTTCGAGTATTTGAAAGCCACCGAGCTGATAGATGATCGATCGAGACTTCTCGG 182  
Qy 148 ProLeuGlyLeuLysTyThrLeuValPheTySerGlyLysAlaProLysGlyThrArg 167  
Db 183 CCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGTTAAAGCCCTTAAGGCACTCGT 242  
Qy 168 ThrSerTrpLeuMetAsnGluTyArgLeuProHisGluThrGluLysTyGlnLys 187  
Db 243 ACTAGTTCGATCATGAACGAGTATCGTCTTCGCAACCATGAACCGAGAGTACCAAAAG 302  
Qy 188 AlaGluLeuSerLeuCyArgValTyLysArgProGlyValGluAspHisProSerVal 207  
Db 303 GCTGAATATCATCTGCGGAGTGTAACAAAGCCAGGAGTAGAAGATCATCCATCGGTA 362  
Qy 208 ProArgSerLeuSerThrArgHisAsnHisAsnSerSerThrSerArgLeuAla 227  
Db 363 CCACGTTCTCTCCACAGACATCATACCATTAATCATCGACATCATCCGTTAGCC 422  
Qy 228 LeuArgGlnGlnGlnHisSerSerSerSerSerSerSerSerSerSerSerSerSerSer 247  
Db 423 TTAAGACAACAACAACACCATTCCTCTCTAATCATTCGACACCAACACCTTAACAAC 482  
Qy 248 AsnAsnAsnLeuAsnAsnLeuLysLeuSerThrGluTySerGlyAspGlySerThr 267  
Db 483 AACACAACATCAACAATCTCGAGAGCTCTCCACCGAATATTCGCGGACGCGAGCACA 542  
Qy 268 ThrThrThrThrThrAsnSerAsnSerAspValThrLeuAlaLeuAlaAsnGlnAsnLeu 287  
Db 543 ACAACAACGACCAACAACAGTAACTCTGAGCTTACCATTCCTAGCCCAATCAACAACA 602  
Qy 288 TyrArgProMetProTyArgProTyArgProTyArgProTyArgProTyArgProTyArg 307  
Db 603 TATCGTCAATGCTTACGACACAGCAACAACATTCATGATGATCTCTACGAGAAATCAT 662  
Qy 308 GlnAspAspGluThrAlaThrLeuValAspLeuGlnArgLeuValAsnTyrGlnLeu 327  
Db 663 CAAGACGATGATGAACATGCTTGTGACGATCTTCAAGAGCTAGTAACTTACCAATA 722  
Qy 328 SerAspGlyAlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetile 347
```

```
Db 723 TCAGATGGAGGACACACGCTAATGCTCAAACTCAAGCGCGTGTAGCTATGACATGATT 782  
Qy 348 ProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleValProAsp 367  
Db 783 CCTCAGAGACGATTCACAAACATGCTTTGTGGGATATGTGGAATCCAATAGTACCAAGAT 842  
Qy 368 GlyAsnArgAspHisIleTyThrAsnIleProPheLys 379  
Db 843 GGAACACAGAGATCACTACTATAATTCCTTTTAAG 878
```

RESULT 4

AKI02902

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone: J033113D13, full

insert sequence.

AKI02902

ACCESSION

AKI02902.1 GI:32988111

VERSION

FLI_CDNA; CAP trapper.

KEYWORDS

Oryza sativa (japonica cultivar-group)

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-Length cDNA Project Team;

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,

Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,

Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,

Ohtsuki, K., Shihiki, T., Foundation of Advancement of International

Genome Sequencing & Analysis Group; Otomo, Y., Murakami, K.,

Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,

Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;

Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,

Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Ota, Y.,

Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

JOURNAL

Science 301 (5631), 376-379 (2003)

PUBMED

12869764

REFERENCE

2

(bases 1 to 1631)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,

Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,

Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,

Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,

Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,

Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,

Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,

Nami, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,

Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,

Oeato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,

Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,

Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,

Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,

Tagawa, A., Takahashi, F., Takaku-Akai, S., Tanaka, T., Tomaru, A.,

Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,

Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and

Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of

Agrobiological Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki

305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,

Tel: 81-29-838-7007, Fax: 81-29-838-7007)

COMMENT This clone is one of the 28k full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Itoh, K., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kuroaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hangaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Osato, N., Oca, Y., Saichou, H., Sakai, C., Sakai, C., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES
 source Location/Qualifiers
 1. 1631
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J033113D13"

ORIGIN

Alignment Scores:
 Pred. No.: 2 45e-54 Length: 1631
 Score: 815.50 Matches: 190
 Percent Similarity: 59.2% Conservative: 49
 Best Local Similarity: 47.0% Mismatches: 97
 Query Match: 40.1% Indels: 69
 DB: 15 Gaps: 10

US-10-780-703-2 (1-379) x AK102902 (1-1631)

QY 42 GluAlaAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGlu 61
 DB 276 GAGGTAGATGGCATGAGGATGACCTGGTGTGATCCCGGGTTCAAGTTCACCGCGGAG 335
 QY 62 GluGluLeuLeuGluPheTyr-LeuArgArgLysValGluGlyLysArgPheAsnValGlu 81
 DB 336 GAGGAGCTGATCGAGTCTACCTCCCGCGGAAGGTGGAGGCAAGCGCTTCAACGTGA 395
 QY 81 uLeuLeuThrPheLeuAspLeuTyrArgTyrAspProThrGluLeuProAlaMetAla 101
 DB 396 GCTCATCACCTCTCGACCTCTACCGCTACGACCCCTGGAGCTCCCGCAATGGCGGC 455
 QY 101 alleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAs 121
 DB 456 GATAGGGGAGAGAGTGGTCTTCTACGTGCTTCGGACAGGAAGTACAGGACGGGGA 515
 QY 121 pArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetI 141
 DB 516 CCGGCGCAACCGGGTACGCGCTCGGGTACTGGAAGGCGACGGGGCGCGCGATGAT 575
 QY 141 eArgSerGluThrSerArgProIleGlyLysThrLeuValPheTyrSerGlyLys 161
 DB 576 CCGAGCCGAGAACACCGCCCCATCGGGCTCAAGAGACGCTTGTCTTACTCCGGCAA 635
 QY 161 sAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisGlu 181

DB 636 GGCCCCCAAGGGCGTCCGACGACGAGTGCATCATGAACGAGTATCGCTTCCCCCGCGCA 695
 QY 181 uThrGluLysTyrGlnLysAlaGluLeuSerLeuLysArgValTyrLysArgProGlyVa 201
 DB 696 CACCGACCGGTACACCAAGACTGAAATCTCACTATGCGCGGTGTACAGCGCACCGGAT 755
 QY 201 lGluAsp---HisProSerVal-ProArg-----SerLeuSerT 213
 DB 756 CGACGACGGCCATGGCCAGGTATCCAGCGGCGGTCTCGCGGCACCTCCCGCGGGCGG 815
 QY 213 hrArgHisHisAsnHisAsnSerThrSerArgLeuAlaLeuArgGln----- 230
 DB 816 CGCGGACCGGTACAGACCAAGCAAGGCTCTCTCGACGCTCCAGCGCGCGCCGCC 875
 QY 231 --GlnGlnHisHisSerSerSerAsnHisSerSerAsnLeuLeuAsnAsnAsnA 250
 DB 876 TCCAATCTCGTCCCAAGCTCCACCTCTCAGCAGCGAGTGCAGCTCGCGCGCCCATG 935
 QY 250 snIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrT 270
 DB 936 CACGGACCGCCCGCCATGGTGGGCGACAGGACCGTCCCGCGGC-----ACCATCA 989
 QY 270 hrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArg- 289
 DB 990 GCAGCAGCAGCAGCTCCAGCTGCCAAGCTTGGCGGCTACTCTGCAGAAATTCCTCGAT 1049
 QY 290 ----ProMetProTyrAspThrSerAsnAsnThrLeuLeuValSerThrArg----- 305
 DB 1050 GGCTTCCGTCCGCGCGCGATCAGCAGCAGCAATTTTCAGCAAGACTTCGCGCGGCATT 1109
 QY 305 ----- 305
 DB 1110 GTACGACGAGTACTCCAGAACAGGAGCGGTGCTTCCGCTCCACGTCCTCGTGTCTCAA 1169
 QY 306 -----AsnHisGlnAspAspGluThrAlaIleValAspAspLeuGlnArgL 322
 DB 1170 CCTCGTCAACGCGGCATCCATG-GGCAGCTCCGCGCTGCAATTGATGAGTGCAGCTCG 1228
 QY 322 euValAsnTyr-----GlnIleSerAspGlyAlaThrThr----- 333
 DB 1229 TGGTGGCCACCGCACCGCTCTTACATTAACCCCGCGCGCGCACCACTATAGTC 1288
 QY 334 -----LeuMetProGlnThrGlnAlaLeuAlaMetAsnMetIleProAlaGly- 350
 DB 1289 AATTCCTTCACTGCTACGCGCTGTCGATCAACCAACGCGCGCGCTCGGGA 1348
 QY 351 -----ThrIlePro-----AsnAsnAla 357
 DB 1349 CGACGACGCGCGCGCGCGCGCTGCGCTGCGCTGCGCGCTTCTCCGACAGGA 1408
 QY 357 euTrpAspMetTrpAsnProIleValProAspGlyAsnArgAspHisTyrThrAsnIleP 377
 DB 1409 TCTGGAGCTGGAAATATCCGATCCCGGCGCGAGGTAGGATTACAGCACCAGTACTG 1468
 QY 377 roPheLys 379
 DB 1469 GATTCAAG 1476
 RESULT 5
 LOCUS AK072275
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone: J023012F04, full insert sequence.
 ACCESSION AK072275
 VERSION AK072275.1 GI:32982298
 KEYWORDS FLI CDNA; CAP trapper.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1
 AUTHORS The Rice Full-Length cDNA Consortium, National Institute of

QY 252 ---AsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAsp----- 264
 Db 912 GCCTCGTGGTCCACCGTCCGAGGAGGAGCGCGTCTTCCACCTCAAG 971
 QY 265 ---GlySerThrThrThrThrThrAsnSerAspValThrLeuAlaLeuAla 283
 Db 972 GCGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1007
 QY 284 AsnGlnAsnIleTyrArgProMetProTyrAspThr-----SerAsnAsn 298
 Db 1008 CACACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1067
 QY 299 ThrLeuIleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAsp 318
 Db 1068 TCAGCAATGCGG-----ACCATCCCATCGAGGAG 1097
 QY 319 LeuGlnArgLeuVal-----AsnTyrGlnIleSer 328
 Db 1098 CTGAGTGGGCGGATGGATGTTACAGCAACATCCACCAAACTTACTCAACT 1157
 QY 329 AspGlyAlaThrThrLeuMetProGlnThrGlnAlaLeuAlaLeuAlaMetAsnMetIlePro 348
 Db 1158 CATCAATCAACTCTTGGCGGCATCGCAACAAAGCCCTCTT-----CTTCCA 1205
 QY 349 AlaGlyThrIleProAsnAlaLeuTyrAspMetTyrAsnProIleVal 365
 Db 1206 TTGCTAGCATG-----GAGAGATTGGGAT---TGGNACCCTCTCTTA 1247

RESULT 6

AK071052

LOCUS

DEFINITION Oriza sativa (japonica cultivar-group) cDNA clone:J023075809, full insert sequence.

ACCESSION

AK071052

VERSION

AK071052.1 GI:32981075

KEYWORDS

FLI_CDNA; CAP trapper.

SOURCE

Oriza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1

The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Negata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Iida, Y., Kuroaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

TITLE

Science 301 (5631), 376-379 (2003)

PUBMED

12869764

REFERENCE

2

(bases 1 to 1891)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Kuroaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ohtsuki, K., Ohtsuki, K., Ohtsuki, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

TITLE

Direct Submission

JOURNAL

Submitted (05-DEC-2001)

Shohei Kikuchi, National Institute of

Agrobiological Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki

305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,

Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica

rice. URL : http://cdna01.dna.affrc.go.jp/cDNA/

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,

Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,

Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Namiki, T.,

Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and

Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,

Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kuroaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,

Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,

Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,

Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center

and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,

Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,

Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,

Ishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,

Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,

Sasa, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,

Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1..1891

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="rRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="J023075E09"

ORIGIN

Alignment Scores:

Pred. No.: 9,066-53 Length: 1891

Score: 796.50 Matches: 177

Percent Similarity: 55.9% Conservative: 45

Best Local Similarity: 44.6% Mismatches: 74

Query Match: 39.2% Indels: 101

DB: 15 Gaps: 14

US-10-780-703-2 (1-379) x AK071052 (1-1891)

QY 44 AspaAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63

Db 364 GACGGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 423

QY 64 LeuIleGluPheTyrLeuArgArgLysValGluGluLysArgPheAsnValGluLeuIle 83

Db 424 CTCATCGAGTCTTACCTCCGCGGAAAGGTGGAGGGGGAAGCGGTTCAACATCGAGCTCATC 483

Qy 84 ThrPheLeuAspLeuTyrArgTyrAspProTTPGluLeuProAlaMetAlaAlaIleGly 103
 Db 484 GCCTTCGTCACCTCTACCGCTATGACCGTGGGATCTCCCGCTTGGCGCTCAATTTGGG 543
 Qy 104 GluLeuGluTyrPheTyrValProArgAspArgLysTyrArgAenGlyAspArgPro 123
 Db 544 GACAGAGGTGGTCTCTATGTGCGAAGGAGCGCAAGTACCGAAGCGCGATCGGCC 603
 Qy 124 AsnArgValThrThrSerGlyTyrTTPLeuAlaThrGlyAlaAspArgMetIleArgSer 143
 Db 604 AACCGAGTACGCCATCGCGGTACTCGAAGGCCACAGCGCGGATAGATGGTGAAGTG 663
 Qy 144 GluThrSerArgProIleGlyLeuLysValThrLeuValPheTyrSerGlyLysAlaPro 163
 Db 664 GAGGTGACCGTCTATTCGGCGCTCAAGAAGACGCTCGTCTTCTACGTGGCGAAGGCC 723
 Qy 164 LysGlyThrArgThrSerTTPLeuMetAsnGluTyrArgLeuProHisHisGluThrGlu 183
 Db 724 AAGGGCTTCGACGACCTGGATCATGAACGATATCGTCTCCCTCATGCGCGCTGAC 783
 Qy 184 LysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGlyValGluAsp 203
 Db 784 CGCTACCAAAAG---GAAATTCGTGTGTCGAGTCTACAAAGCGCCAGCAATCGAAGAC 840
 Qy 204 His-----ProSerValProArgSerLeu 211
 Db 841 AACTTCCACCTCCCGGCACCAACCACTAAATCTCCGGCTCAAGAGCGCGCGCGGCG 900
 Qy 212 SerThrArgHis-----HisAsnHisAsnSerThrSerArgLeuAlaLeuArg 229
 Db 901 GGCAGAGACCGCGCGCGGACGCGACGTCATCGACGGGGGGTGGCGGCGCGCGC 960
 Qy 230 GlnGlnGln-----HisHisSerSer----- 236
 Db 961 CTCGCGCGCGGTTCGACGGCGCGCGCGCGCGCCACCTCATCGCGCTCATGGGAAG 1020
 Qy 237 ---SerSerAsnHisSerAspAsnLeuAsnAsnAsnLeu----- 251
 Db 1021 GCAACATACACGATCATCACACCGCGCGCGCGAGGAGCGGACGCGCATCTCACTG 1080
 Qy 251 ----- 251
 Db 1081 TCGCGCGCGCGCGGTGCTGATCTTCGCGCGCGCGCGGTTCGCGTGCAGCGCGTGGTG 1140
 Qy 252 ---AsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAsp----- 264
 Db 1141 GCCTCGGTGAGCTCCACCGCTGCGACGAGGAGGAGCGCGCGCTCTCCACCTCAAG 1200
 Qy 265 ---GlySerThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAla 283
 Db 1201 GCGCGCGCGAGCGAGCGAAGACGACG-----ACGATGCAT 1236
 Qy 284 AsnGlnAsnIleTyrArgProMetProTyrAspThr-----SerAsnAsn 298
 Db 1237 CATCACCATCATCATCAGCAGCTGCTTCAACTCATGCTCTGCTCAACACCGCATCG 1296
 Qy 299 ThrLeuIleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAsp 318
 Db 1297 TCAGCAATGGC-----ACCATCCCATCGACGAG 1326
 Qy 319 LeuGlnArgLeuVal-----AsnTyrGlnIleSer 328
 Db 1327 CTGAGTAGGGCGATTTGGATCTTACAGCAACATCCACCATCAACCAAACTCTACTACAACT 1386
 Qy 329 AspGlyAlaThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIlePro 348
 Db 1387 CATCAATCAACTGTTTGGCGCCCATCGCAAAAGCCCTCTT-----CTTCCA 1434
 Qy 349 AlaGlyThrLeuProAsnAsnAlaLeuTyrAspMetTyrAsnProIleVal 365
 Db 1435 TTCCTAGCATG-----GAGAAGATTGGAT---TGGAACTCTCTCTTA 1476

RESULT 7

AK068393
 LOCUS
 DEFINITION

AK068393 2424 bp mRNA linear PLN 24-JUL-2003
 Oryza sativa (japonica cultivar-group) cDNA clone:J013151109, full
 insert sequence.

ACCESSION

AK068393.1 GI:32978411

VERSION

FLI CDNA; CAP trapper.

KEYWORDS

Oryza sativa (japonica cultivar-group)

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 The Rice Full-Length cDNA Consortium, National Institute of
 Agronomical Sciences Rice Full-Length cDNA Project Team:
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group: Ohtsuki, K., Murakami, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, K., Ota, Y.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice

JOURNAL

Science 301 (5631), 376-379 (2003)

PUBLISHED

12869764

REFERENCE

2 (bases 1 to 2424)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T.,
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
 Nami, T., Nariakawa, R., Niihara, J., Nishi, K., Nomura, K.,
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
 Osato, N., Ota, Y., Ohtsuki, T., Saitoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sugabe, Y., Sugano, S.,
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
 Yoshimura, A.

TITLE

JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
 Agronomical Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28K full-length cDNA clones from japonica
 rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cdna/>
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T.,
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and
 Yamamoto, M.
 PAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ootani, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. 2424
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013151.09"

ORIGIN

Alignment Scores:
Pred. No.: 4,378-52 Length: 2424
Score: 789.50 Matches: 175
Percent Similarity: 55.7% Conservative: 46
Best Local Similarity: 44.1% Mismatches: 75
Query Match: 38.8% Indels: 101
DB: 15 Gaps: 14

US-10-780-703-2 (1-379) x AK068393 (1-2424)

QY 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63
DB 135 GACGGCGAGCGGACACATGATGTCGGGGTTCGGGTTCCACCGGAGGAGGAG 194
QY 64 LeuileGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuile 83
DB 195 CTCATCGAGTTCTACCTCCGGGAAAGTGGAGGGAAGCGGTTCAACATCGAGTCTATC 254
QY 84 ThrPheLeuAspLeuThrArgThrArgProThrGluLeuProAlaMetAlaLeuileGly 103
DB 255 GCCTTCGTCGACCTCTACCGCTATGACCGTGGATCTTCCCGCTCTGGCCTCAATTGGG 314
QY 104 GluLysGluThrPheThrValProArgAspArgLysThrArgAsnGlyAspArgPro 123
DB 315 GACAGAGTGGTTCTTCTATGTCCAGGAGCCGCAAGTCCGAAACGGCGATCGGCC 374
QY 124 AsnArgValThrThrSerGlyThrTrpLysAlaThrGlyAlaAspArgMetIleArgSer 143
DB 375 AACCGAGTACCGCATCGGGTACTGGAAGGCCACAGGGCGGATAGGTGTGAAGTGG 434
QY 144 GluThrSerArgProLleGlyLysLysThrLeuValPheThrLysArgGlyLysAlaPro 163
DB 435 GAGGGTGACCGTCTATCGGCTCAAGAAGCGCTCTTCTTACGTCGCAAGGCGGCC 494
QY 164 LysGlyThrArgThrSerTrpIleMetAsnGluThrArgLeuProHisGluThrGlu 183
DB 495 AAGGGGCTTCGCGAGCGTGGATCATGAACGAGTATCGTCTCCCTCATGGCGCGTGCAC 554
QY 184 LysThrGlnLysAlaGlaGlaSerLeuCysArgValThrLysArgProGlyValGluAsp 203
DB 555 CGCTACCAAG---GAAATTTCGTGTCTGAGTCTACAAACGCCCGAGATCGAGAC 611
QY 204 His-----ProSerValProArgSerLeu 211
DB 612 AACTTCCACCTCAGCGGCACACCACTAAATCTCCGGCTCAAGGCGCGCGCGATG 671

QY 212 SerThrArgHis-----HisAsnHisAsnSerSerThrSerArgLeuAlaLeuArg 229
DB 672 GCGAAGAAGACGCGCGCGGAAACCGGAGTCTATCGAGCGCGGCTCGCGCGCGCCG 731
QY 230 GlnGlnGln-----HisHis-----Ser 235
DB 732 CTCGCCCGATGTTTCGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 791
QY 236 SerSerSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsnAsnAsn 251
DB 792 GCAACATACAGCATCATCACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 851
QY 251 ----- 251
DB 852 TCG 911
QY 252 ---AsnAsnLeuGluLysLeuSerThrGluThrArgGlyAsp----- 264
DB 912 GCCTCGCTGAGCTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 971
QY 265 ---GlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 283
DB 972 GCG 1007
QY 284 AsnGlnAsnIleTyrArgProMetProTyrAspThr-----SerAsnAsn 298
DB 1008 CATCACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1067
QY 299 ThrLeuLeuValSerThrArgAsnHisGlnAspAspGluThrAlaLeuValAspAsp 318
DB 1068 TCAGCAATGGCG-----ACCATCCCATCATCATCATCATCATCATCATCATCAT 1097
QY 319 LeuGlnArgLeuVal-----AsnTyrGlnIleSer 328
DB 1098 CTGAGTAGGGGATTGGATCTTACAGCAACATCCACCATCAACCAACTCTACTACA 1157
QY 329 AspGlyAlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIlePro 348
DB 1158 CATCAATCACTGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1205
QY 349 AlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleVal 365
DB 1206 TTCGCTAGCATG-----GAGAAGATTGGGAT---TGAACCTCTCTCCTA 1247
RESULT 8
AY672069 1406 bp mRNA linear PLN 25-SEP-2004
LOCUS Hordeum vulgare subsp. vulgare NAC transcription factor mRNA,
DEFINITION complete cds.
ACCESSION AY672069
VERSION AY672069.1 GI:52352766
KEYWORDS
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1. (bases 1 to 1406)
AUTHORS Robertson, M.
TITLE Two Transcription Factors Are Negative Regulators of Gibberellin
Response in the HvSPY-Signaling Pathway in Barley Aleurone
Plant Physiol. 136 (1), 2747-2761 (2004)
JOURNAL
PUBMED 15347799
REFERENCE 2 (bases 1 to 1406)
AUTHORS Robertson, M.
TITLE Direct Substitution
JOURNAL Submitted (29-JUN-2004) Plant Industry, CSIRO, GPO Box 1600,
Canberra, ACT 2601, Australia
FEATURES
Location/Qualifiers
1. 1406
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Himalaya"


```

/ sub_species="vulgare"
/ db_xref="taxon:112509"
4..1179
/ note="HSINAC; HvSPY-interacting NAC"
/ codon_start=1
/ product="NAC transcription factor"
/ protein_id="AAU43824.1"
/ db_xref="GI:52352767"
/ translation="MTRSRPTTNGSVDPDHHQQOQHDGVGGGLQHGGERVTVM
PGRFPHYRDEELIEFYLRKVDGKRFNDLIASVDLYRDPDLPALASIGDKWEFFY
VPRDRYRNDNRNPTSPGKATGADRMVKVVEGNSIGLKTTLVYVYGKAPGLR
SSWIMNEYRPHQPTERYQKEISLCRVYKRPDDNLHLGTLTTRSSGSRARTRST
AAHTSAVAHQPOPEVEGAGHSSALKAYNAHTTQGTNAASTWTLSSAAGVTPP
AMERSASVASLSTSTSEEDCTSLYHHHLKGNHNPVLMOLHSSTHAALLNTNSSAMA
TIPIDELSRAGISYSQASNPENQPTAPMQGPLLNPSPLEKLDWNPLESFKVCTSPK"

ORIGIN
Alignment Scores:
Pred. No.: 7,1e-52 Length: 1406
Score: 783.00 Matches: 181
Percent Similarity: 57.3% Conservative: 54
Best Local Similarity: 44.1% Mismatches: 82
Query Match: 38.5% Indels: 93
DB: 15 Gaps: 15

US-10-780-703-2 (1-379) x AY672069 (1-1406)

QY 21 GluLysGlyLeuGluAspAsnAspHisArgGlyGly-----Gln-GluSerHisVa 37
DB 3 GATGACTCGATCAAGGCGGACGACCATGGGGGATGATGCGCCAGCAACACCA 62
QY 37 lGlnAsnGluaspGluAlaasp-----AspHis-- 48
DB 63 GCAGCAGCAGCAGCAGCGGGAAGTGGACGCTGCAGCGCGGCGGACGACGT 122
QY 49 -AspMetValMetProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTy 68
DB 123 GGAGACGGTGATGCCCGGGTTCGTTTCCACCCGACGAGGAGGAGCTGATCGAGTTCTA 182
QY 68 rLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLe 88
DB 183 CCTTCGTCGACGAGCGGACGCGGACGCTTCAACATCGACTCATCGCTCCGTCGACCT 242
QY 88 uTyArgTyArgProThrGluLeuProAlaMetAlaAlaLeuGlyGlyGluThrTy 108
DB 243 CTACCGCTACGACCATGGGATCTCCCGCAGTACGCTGCGGACGAGGAGTGTT 302
QY 108 rPheTyValProArgAspArgLysTyArgAsnGlyAspArgProAsnArgValThrTh 128
DB 303 CTTCTAGCTTCTCGGACCGCAAGTACCGGACCGGACCGGACCGGACGAGGAGTGACGCC 362
QY 128 rSerGlyTyTrpLysAlaThrGlyAlaAspArgMetIleArg---SerGluThrSerAr 147
DB 363 GTCAGGGTACTGGAAGCAACAGGGGCGGACGAGTGTGAAAGTCGTGGAGGCAACCG 422
QY 147 gProIleGlyLeuLysLysThrLeuValPheTySerGlyLysAlaProLysGlyThrAr 167
DB 423 CTCCATCGGCTCAAGAAGACGCTCGCTTCTACGTTGGCAAGGACCCCAAGGAGCTCCG 482
QY 167 gThrSerTrpIleMetAsnGluTyArgLeuProHisHiscLurThrGluLysTyTrGlnLy 187
DB 483 CAGCAGCTGATCATGAAACGAGTACCGCTCCGACACGCGGTGAACACCGAGCTACCAAA 542
QY 187 sAlaGluLeuSerLeuCyArgValTyLysArgProGlyValGluAspHis----- 204
DB 543 G---GAATTTTCGCTCTCGCGGGTCTATAACGCCCGGAGGATTCAGCACAACCTTCACCT 599
QY 205 -----ProSerValProArgSerLeuSerThrArgHisHisAsnHi 218
DB 600 CACCGGCACACACAAAGGTCGTCTGCTCGAGCGCGGCGGACGAGGACAGTACAGC 659
QY 218 sAsnSerSerThrSerArgLeuAlaLeuArgGlnGln----- 232

```


QY 52 MetProGlyPheArgPheHisProThrGluGluLeuLeuLeuGluPheTyrLeuArgArg 71
Db 1 ATGCCCGGGTTACAGTTCCACCCGACGAGGAGAGCTGATCGAGTTCTACCTCCGCGG 60
QY 72 LysValGluGlyLysArgPheAsnValGluLeuLeuLeuPheLeuAspLeuTyrArgTyr 91
Db 61 AAGGTGGAGGGCAAGCGCTTCAACGTCGAGCTCATCACCCTTCCTCGACCTCTACCGCTAC 120
QY 92 AspProTyrGluLeuPro-----Ala 98
Db 121 GACCCCTGGAGCTCCCGGAATTGAGCATCTAATTAAACAAGCTCGATGGGGGTGACGCA 180
QY 99 MetAlaAlaGlyGlyLysGluTyrPheTyrValProArgAspArgLysTyrArg 118
Db 181 ATGGCGCGCATAGGGGAGAGAGTGTCTTCTACGTGCTCGGACAGGAAGTACAGG 240
QY 119 AsnGlyAspArgProAsnArgValThrSerGlyTyrTriPylsAlaThrGlyAlaAsp 138
Db 241 AACGGGACCGCGCAACCGGTGACGGGTCTGGGTACTGGAAGCGGACGGGGCGGAC 300
QY 139 ArgMetIleArgSerGluThrSerArgProIleGlyLeuLysValThrLeuValPheTyr 158
Db 301 CGGATGATCCGAGCGAGAACACCGCCCATCGGGCTCAAGAAGACGCTTGTCTCTAC 360
QY 159 SerGlyLeuAlaProLysGlyThrArgThrSerTriPylMetAsnGluTyrArgLeuPro 178
Db 361 TCGGCAAGCGCCCAAGGGCGTCCGACGAGCTGGATCATGAACGAGTATCGCTTCC 420
QY 179 HisHisGluThrGluLysTyrGlnLysAlaGluLeuSerLeuCysArgValTyrLysArg 198
Db 421 CCGCGGACACCGCGCTACCAACAGACTGAAATCTCATATGCGCGGTGTACAGCGC 480
QY 199 ProGlyValGluAsp---HisProSerVal-ProArg-----Se 210
Db 481 ACCGATCATGACGACGCGCATGCGCAGGTATCCAGCGCGGTGTCGCGCGCATCCCGC 540
QY 210 rLeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgG1 230
Db 541 GCGCGCGCGCGCGCGGTACGACACAGCAAGGCTCTCTCGTCCGAGCTCCAGCGCG 600
QY 230 n-----GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 601 AGCGCGCTCCCAACTCCGTCACGCTCCACCTCTCAGCAGCGGTCAGCGCGCGCGCC 660
QY 247 nAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerTh 267
Db 661 GCCATGTCACGACGACCGCGCCATGTTGGGCGCAACAGCGCGCTGCGCGCGCGC--- 715
QY 267 rThrThrThrThrThrAsnSerSerSerSerSerSerSerSerSerSerSerSerSer 287
Db 716 -ACCATCAGCAGCAGCAGCAGCTCCAGCTGCGCAAGCTTGGCGGGCTACCTGCGAGAT 774
QY 287 eTyrArg-----ProMetProTyrAspThrSerAsnAsnThrLeuLeuValSerThrAr 305
Db 775 TCTCGATGCTTCCGCTCGCGCGCGCGATCAGCAGCAGCAATTTCCAGCAAGACTTCGG 834
QY 305 g----- 305
Db 835 GCGGCAATTGACACGAGTACTCCAGAACACGAGCGGTGCTTCGCTCCACGTAAGTACG 894
QY 306 -----AsnHisGlnAspAspAspGluThrAlaLeuValAspAspLe 319
Db 895 CTGCTCAACCTGTCACGCGCATCCATG-GGCAGCTCCGCGGTGCGCATTTGATGAGCT 953
QY 319 uGlnArgLeuValAsnTyr-----GlnIleSerAspGlyAlaThr----- 333
Db 954 GAGCTCGCTGTGGGCGCACGCGCGCTTACATTAACCCCGCGCGCGCGCGCACAA 1013
QY 334 -----LeuMetProGlnThrGlnAlaLeuAlaMetAsnMetIlePr 348
Db 1014 CTATAGTCAATTCCTTCACTGCTACGACGCGCGTCTGCGCATCAACCAACCGCGCGC 1073
QY 348 oAlaGly-----ThrIlePro-----As 354

Db 1074 GCTCGGACGACGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1133
QY 354 nAsnAlaLeuTyrAspMetTyrAsnProIleValProAspGlyAsnArgAspHisTyrTh 374
Db 1134 CGACAGGATCTGGGACTGGGAATAATCCGATCCCGAGGCGGAGGTAGGATTACAGCAC 1193
QY 374 rAsnIleProPheLys 379
Db 1194 CAGTACTGGATTCAG 1209
RESULT 10
AP007649 80577 bp DNA linear HTG 28-DEC-2004
LOCUS
DEFINITION
Lotus corniculatus var. japonicus chromosome 4 clone LjT35A11, ***
SEQUENCING IN PROGRESS ***, 15 unordered pieces.
ACCESSION
AP007649.1 GI:56805960
VERSION
HTG; HTGS PHASE1.
KEYWORDS
Lotus corniculatus var. japonicus (Lotus japonicus)
SOURCE
Lotus corniculatus var. japonicus
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
1
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence
Features and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
2 (bases 1 to 80577)
Sato, S.
Direct Submission
Submitted (28-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex.2337),
Fax: 81-438-52-3934)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1062: contig of 1062 bp in length
* 1063 1162: gap of unknown length
* 1163 2101: contig of 939 bp in length
* 2102 2201: gap of unknown length
* 2202 3130: contig of 929 bp in length
* 3131 3230: gap of unknown length
* 3231 4336: contig of 1106 bp in length
* 4337 4436: gap of unknown length
* 4437 6409: contig of 1973 bp in length
* 6410 6509: gap of unknown length
* 6510 8323: contig of 1814 bp in length
* 8324 8423: gap of unknown length
* 8424 11366: contig of 2943 bp in length
* 11367 11466: gap of unknown length
* 11467 14483: contig of 3017 bp in length
* 14484 14583: gap of unknown length
* 14584 18560: contig of 3977 bp in length
* 18561 18660: gap of unknown length
* 18661 25852: contig of 7192 bp in length
* 25853 25952: gap of unknown length
* 25953 33903: contig of 7951 bp in length
* 33904 34003: gap of unknown length
* 34004 42262: contig of 8259 bp in length
* 42263 43363: contig of 11281 bp in length
* 43364 53743: gap of unknown length
* 53744 70427: contig of 16684 bp in length

Db 37069 CTGGAACATATTATGATTACACGCATGAATTGTGAATTCTGACACTCATGCTCTGTGAC 37128
 QY 186 ----- 186

Db 37129 TTTGATCAATGATGATAGTACTAGACATTTCTCAATTTCTATCTCTAGTTACATATATTC 37188
 QY 187 ----- LysAlaGluIleSerLeuCY 193
 Db 37189 ATTGATTGATTGAATTTGAATTTGGATCAATTTGAAGCTGAGATATCGCTTTG 37248

QY 193 sArgValTyrLysArgProGlyValGluAspHisProSerValProArgSerLeuSerTh 213
 Db 37249 CCGGGTATCAAGAGAGCTGGAGTAGAGATCATCTCTCCCTCGTGTGTCCTCAAT 37308

QY 213 rArgHisHisAsnSerSerThrSerArgLeuAlaLeuArgGlnGlnHis 233
 Db 37309 TACCAAGCAGCCATCATCTCAAGATCATCAAAACCAAGTGAATAGCTCAGCATCAATA 37368

QY 233 s-----HisSerSerSerSerAsnHisSerAspAs 243
 Db 37369 TCTGCCACACATCGGCTTATTCTTGGCCCAATCAATTCAAACCAAGTTCATGATGG 37428

QY 243 nAsnLeuAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerG1 263
 Db 37429 ACAATAGTTGAGAT-----CATCAACGGAGGGAATTAGCAGCACTGTTCCGG 37479

QY 263 yAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 283
 Db 37480 TTCGGGTTTCAGAACATCATGTAAACACA-----GCTCTTGGGCTCTC 37521

QY 283 aAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsn-----AsnThr----- 299
 Db 37522 CAAGTACATAGTTACCGTGTGCGGTGGTTCATCAGCAACACAAATACATAGGACT 37581

QY 300 -----LeuIleValSerThrArgAsnHisGln----- 308
 Db 37582 GCCAGCGGAGGAGGAGGTTGATGCTCTCGCGCAGCACCACCAATCTAATAATCCAAA 37641

QY 308 ----- 308

Db 37642 GCAACACCAATAACTCTTTAGTCCATGTCATGTTTCCCTGCAAAATGCTGTCCTTCAA 37701

QY 309 ----AspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGln-- 326
 Db 37702 TAATAACAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 37761

QY 327 -----IleSerAspGlyAlaThrThrLeuMetProGlnTh 338
 Db 37762 ACAGTTCTTCAATGTTCAAGTCATCCCAATCATCAATTTATCTGCTTTACTGATGCAAA 37821

QY 338 rGlnAlaLeuAlaLeuAlaMetAsnMetIleProAlaGlyThrIlePro-----AsnAs 355
 Db 37822 GCGGCCCAATGTTCTCACTACTACTACCA---AATACACTTCCAACTACCTCTCTGA 37878

QY 355 nAlaLeuTrpAspMetTrpAsnProIleValProAspGlyAsnArgAspHisTyrThrAs 375
 Db 37879 CCGATCTGGGAC---TGAATCCATC---CTGAGGTAAACCAAGACCACCCCAACAT 37932

QY 375 nIleProPhelys 379
 Db 37933 CATGCCCTTCAAG 37945

RESULT 11
 LOCUS AP007409
 DEFINITION Lotus corniculatus var. japonicus chromosome 4 clone LjT46D13, ***
 SEQUENCING IN PROGRESS ***, 24 unordered pieces.
 ACCESSION AP007409
 VERSION AP007409.1 GI:56805724
 KEYWORDS HTG; HTGS PHASE1.
 SOURCE Lotus 'corniculatus' var. japonicus
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
 1
 Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 Structural Analysis of a Lotus japonicus Genome. XI. Sequence Features and Mapping of Nine hundred twenty-one TAC Clones Unpublished
 2 (bases 1 to 91718)
 Sato, S.
 Direct Submission
 Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba, 292-0878, Japan (E-mail: ssato@kazusa.or.jp).
 URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex.2337), Fax: 81-438-52-3934
 * NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 3414: contig of 3414 bp in length
 * 3415 3514: gap of unknown length
 * 3515 4430: contig of 916 bp in length
 * 4431 4531: gap of unknown length
 * 4531 6016: contig of 1486 bp in length
 * 6017 6116: gap of unknown length
 * 6117 7544: contig of 1428 bp in length
 * 7545 7644: gap of unknown length
 * 7645 8647: contig of 1003 bp in length
 * 8648 8747: gap of unknown length
 * 8748 10035: contig of 1288 bp in length
 * 10036 10135: gap of unknown length
 * 10136 11908: contig of 1773 bp in length
 * 11909 12008: gap of unknown length
 * 12009 13497: contig of 1489 bp in length
 * 13498 13597: gap of unknown length
 * 13598 15654: contig of 2057 bp in length
 * 15655 15754: gap of unknown length
 * 15755 18265: contig of 2511 bp in length
 * 18266 18365: gap of unknown length
 * 18366 20730: contig of 2365 bp in length
 * 20731 20830: gap of unknown length
 * 20831 22099: contig of 1269 bp in length
 * 22100 22199: gap of unknown length
 * 22200 25367: contig of 3168 bp in length
 * 25368 25467: gap of unknown length
 * 25468 29301: contig of 3834 bp in length
 * 29302 29401: gap of unknown length
 * 29402 32762: contig of 3361 bp in length
 * 32763 32862: gap of unknown length
 * 32863 34190: contig of 1328 bp in length
 * 34191 34290: gap of unknown length
 * 34291 38886: contig of 4596 bp in length
 * 38887 38986: gap of unknown length
 * 38987 43709: contig of 4723 bp in length
 * 43710 43809: gap of unknown length
 * 43810 47835: contig of 4026 bp in length
 * 47836 47935: gap of unknown length
 * 47936 54444: contig of 6509 bp in length
 * 54445 54544: gap of unknown length
 * 54545 60454: contig of 5910 bp in length
 * 60455 60554: gap of unknown length
 * 60555 70888: contig of 10334 bp in length
 * 70889 70988: gap of unknown length
 * 70989 88088: contig of 17100 bp in length
 * 88089 88188: gap of unknown length
 * 88189 91718: contig of 3530 bp in length.

Location/Qualifiers
 1. .91718

FEATURES
 source

QY 186 ----- 186
 Db 41929 TTTAAGTAGCAAAATTTAGATCCAAAGCTAGCTATAGCATGCCAACACAGTAGGTGGTG 41988
 QY 186 ----- 186
 Db 41989 AACTCGCAGCCACCATTAAATTTATTTCAAGAGTAGTCTTTTGGCATTTAATAAATACT 42048
 QY 186 ----- 186
 Db 42049 GATAGAAAAGATTTTCAGGAATGTTTCAGCGTTGTAGCACCGTTTGAATAATTAACCTCA 42108
 QY 186 ----- 186
 Db 42109 CTCGAACATATTATGATTATACGCAAGAATTGTGAATTCGACATCATGTCTCCGGAC 42168
 QY 186 ----- 186
 Db 42169 TTTGATCAATGATGATAGTCTAGACATTTCTCAATTTCTAGTTACATATATTC 42228
 QY 187 ----- 187
 Db 42229 ATTGATTGATTGAATGTGAATGTGAATTTGGATCAATAGAGGCTGAGATATCGTTTG 42289
 QY 193 sargValTyLysArgProGlyValGluAspHisProSerValProArgSerLeuSerTh 213
 Db 42289 CCGGTATACAGAGAGCTGGAGTAGAGATCATCTTCCCTCCCTCGTTGTCTCCCAAT 42348
 QY 213 rArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHis 233
 Db 42349 TACCAAGCAGCCATCATCTCAAGATCATCAAAACCAAGTGAATAGCTCAGCATCAATA 42408
 QY 233 s-----HisSerSerSerSerAsnHisSerAspAs 243
 Db 42409 TCTGCCACACATCGGGCTTATTCTTGGCCCAATCAATCAAAACCAAGTTCATGATGG 42468
 QY 243 nAsnLeuAsnAsnAsnAsnAsnLeuLeuLeuLeuLeuLeuSerThrGluTySerG 263
 Db 42469 ACAATAGTTGAGAT-----CATGAACGGAGGAGGATTTAGCAGCACTGGTTCGGG 42519
 QY 263 yAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 283
 Db 42520 TTCGGGTTTCAAGATCATGTATCAACA-----GCTCTGGGCTCTC 42561
 QY 283 aAsnGlnAsnLeuTyArgProMetProTyArgProMetProTyArgProMetProTyArg 299
 Db 42562 CAATGATCAATGCTTACCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 42621
 QY 300 -----LeuLeuValSerThrArgAsnHisGln----- 308
 Db 42622 GCCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 42681
 QY 308 ----- 308
 Db 42682 GCAACCAACCAATACTTTTAGTCCATGCTTTTCCCTGCAATGCTGTGCTTCAAA 42741
 QY 309 -----AspAspAspGluThrAlaAlaLeuValAspAspLeuGlnValAsnTyGln-- 326
 Db 42742 TAATAACAATAATATGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 42801
 QY 327 -----IleSerAspGlyAlaThrThrLeuMetProGlnTh 338
 Db 42802 ACAGTTCTCAATGTTCAAGTCAATCCCAATCATCAATTTATCTGTTTACTGATGCAAC 42861
 QY 338 rGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThrIlePro-----AsnAs 355
 Db 42862 GCGCCACCAATGTCTCACTCACTACCA--AATACATCTTCAACTACCTTCTCTGA 42918
 QY 355 nAlaLeuTrpAspMetTrpAsnProIleValProAspGlyAsnArgAspHisTyThrAs 375
 Db 42919 CCGACTGTGGGAC---TGGAATCCAATC---CTGAGGTAAACCAAGACCAACCAACAT 42972
 QY 375 nIleProPheTyS 379

Db 42973 CATGCCCTTCAAG 42985
 RESULT 12
 LOCUS AX654570
 DEFINITION Sequence 4440 from Patent WO03000898.
 ACCESSION AX654570
 VERSION AX654570.1 GI:29157384
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
 TITLE Plant genes involved in defense against pathogens
 JOURNAL Patent: WO '03000898-A 4440 03-JAN-2003;
 Syngenta Participations AG (CH)
 FEATURES
 source
 1. 1250
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,638-50 Length: 1250
 Score: 759.00 Matches: 174
 Percent Similarity: 50.8% Conservative: 45
 Best Local Similarity: 40.4% Mismatches: 72
 Query Match: 37.3% Indels: 140
 DB: Gaps: 14
 US-10-780-703-2 (1-379) x AX654570 (1-1250)
 QY 50 MetValMetProGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyLeu 69
 Db 1 ATGGTAATGCGGGGTTCGGTTCACCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
 QY 70 ArgArgLysValGluGlyArgPheAsnValGluLeuIleThrPheLeuAspLeuTy 89
 Db 61 CGGCAAGAGTGGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 QY 90 ArgTyArgProTrpGluLeuProAlaMetAlaAlaGlyGlyGlyGlyGlyGlyGlyGlyGly 109
 Db 121 CGCTATGACCGTGGGATCTTCCGCTCTGGCTCAATTGGGACAAAGGAGTGGTCTTCTTC 180
 QY 110 TyrValProArgAspArgLysTyArgAsnGlyAspArgProAsnArgValThrThrSer 129
 Db 181 TATGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 QY 130 GlyTyTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIle 149
 Db 241 GGGTACTGGAGGCCACAGGGCGGATAGGATGGTGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 QY 150 GlyLeuLysLysThrLeuValPheTySerGlyLysAlaProLysGlyThrArgThrSer 169
 Db 301 GGCCTCAAGAAGACGCTCGTCTTCTACGTGGCAAGGCGCCCAAGGGGCTTCGACGACG 360
 QY 170 TrpIleMetAsnGluTyArgLeuProHisHisGluThrGluLysTyGlnLys----- 187
 Db 361 TGGATCATGAACAGATATCGTCTCCCTCATGCGACGCTGACCGCTACCAAAAGATAGCT 420
 QY 187 ----- 187
 Db 421 TTTGATATAGTACCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
 QY 188 -----AlaGlu 189
 Db 481 TGTCTCTCAGCTGGTGATGAAATTAAAGTACACACCATCGGAAATTTGTTATGTAGAGGAA 540

AP008207_021	2100001	2210000	AP008207_094	9400001	9510000
AP008207_022	2200001	2310000	AP008207_095	9500001	9610000
AP008207_023	2300001	2400000	AP008207_096	9600001	9710000
AP008207_024	2400001	2510000	AP008207_097	9700001	9810000
AP008207_025	2500001	2610000	AP008207_098	9800001	9910000
AP008207_026	2600001	2710000	AP008207_099	9900001	1001000
AP008207_027	2700001	2810000	AP008207_100	10000001	1011000
AP008207_028	2800001	2910000	AP008207_101	10100001	1021000
AP008207_029	2900001	3010000	AP008207_102	10200001	1031000
AP008207_030	3000001	3110000	AP008207_103	10300001	1041000
AP008207_031	3100001	3210000	AP008207_104	10400001	1051000
AP008207_032	3200001	3310000	AP008207_105	10500001	1061000
AP008207_033	3300001	3410000	AP008207_106	10600001	1071000
AP008207_034	3400001	3510000	AP008207_107	10700001	1081000
AP008207_035	3500001	3610000	AP008207_108	10800001	1091000
AP008207_036	3600001	3710000	AP008207_109	10900001	1101000
AP008207_037	3700001	3810000	AP008207_110	11000001	1111000
AP008207_038	3800001	3910000	AP008207_111	11100001	1121000
AP008207_039	3900001	4010000	AP008207_112	11200001	1131000
AP008207_040	4000001	4110000	AP008207_113	11300001	1141000
AP008207_041	4100001	4210000	AP008207_114	11400001	1151000
AP008207_042	4200001	4310000	AP008207_115	11500001	1161000
AP008207_043	4300001	4410000	AP008207_116	11600001	1171000
AP008207_044	4400001	4510000	AP008207_117	11700001	1181000
AP008207_045	4500001	4610000	AP008207_118	11800001	1191000
AP008207_046	4600001	4710000	AP008207_119	11900001	1201000
AP008207_047	4700001	4810000	AP008207_120	12000001	1211000
AP008207_048	4800001	4910000	AP008207_121	12100001	1221000
AP008207_049	4900001	5010000	AP008207_122	12200001	1231000
AP008207_050	5000001	5110000	AP008207_123	12300001	1241000
AP008207_051	5100001	5210000	AP008207_124	12400001	1251000
AP008207_052	5200001	5310000	AP008207_125	12500001	1261000
AP008207_053	5300001	5410000	AP008207_126	12600001	1271000
AP008207_054	5400001	5510000	AP008207_127	12700001	1281000
AP008207_055	5500001	5610000	AP008207_128	12800001	1291000
AP008207_056	5600001	5710000	AP008207_129	12900001	1301000
AP008207_057	5700001	5810000	AP008207_130	13000001	1311000
AP008207_058	5800001	5910000	AP008207_131	13100001	1321000
AP008207_059	5900001	6010000	AP008207_132	13200001	1331000
AP008207_060	6000001	6110000	AP008207_133	13300001	1341000
AP008207_061	6100001	6210000	AP008207_134	13400001	1351000
AP008207_062	6200001	6310000	AP008207_135	13500001	1361000
AP008207_063	6300001	6410000	AP008207_136	13600001	1371000
AP008207_064	6400001	6510000	AP008207_137	13700001	1381000
AP008207_065	6500001	6610000	AP008207_138	13800001	1391000
AP008207_066	6600001	6710000	AP008207_139	13900001	1401000
AP008207_067	6700001	6810000	AP008207_140	14000001	1411000
AP008207_068	6800001	6910000	AP008207_141	14100001	1421000
AP008207_069	6900001	7010000	AP008207_142	14200001	1431000
AP008207_070	7000001	7110000	AP008207_143	14300001	1441000
AP008207_071	7100001	7210000	AP008207_144	14400001	1451000
AP008207_072	7200001	7310000	AP008207_145	14500001	1461000
AP008207_073	7300001	7410000	AP008207_146	14600001	1471000
AP008207_074	7400001	7510000	AP008207_147	14700001	1481000
AP008207_075	7500001	7610000	AP008207_148	14800001	1491000
AP008207_076	7600001	7710000	AP008207_149	14900001	1501000
AP008207_077	7700001	7810000	AP008207_150	15000001	1511000
AP008207_078	7800001	7910000	AP008207_151	15100001	1521000
AP008207_079	7900001	8010000	AP008207_152	15200001	1531000
AP008207_080	8000001	8110000	AP008207_153	15300001	1541000
AP008207_081	8100001	8210000	AP008207_154	15400001	1551000
AP008207_082	8200001	8310000	AP008207_155	15500001	1561000
AP008207_083	8300001	8410000	AP008207_156	15600001	1571000
AP008207_084	8400001	8510000	AP008207_157	15700001	1581000
AP008207_085	8500001	8610000	AP008207_158	15800001	1591000
AP008207_086	8600001	8710000	AP008207_159	15900001	1601000
AP008207_087	8700001	8810000	AP008207_160	16000001	1611000
AP008207_088	8800001	8910000	AP008207_161	16100001	1621000
AP008207_089	8900001	9010000	AP008207_162	16200001	1631000
AP008207_090	9000001	9110000	AP008207_163	16300001	1641000
AP008207_091	9100001	9210000	AP008207_164	16400001	1651000
AP008207_092	9200001	9310000	AP008207_165	16500001	1661000
AP008207_093	9300001	9410000	AP008207_166	16600001	1671000

AP008207_126	12600001	12710000	Pred. No.:	7.67e-41	Length:	110000
AP008207_127	12700001	12810000	Score:	672.50	Matches:	189
AP008207_128	12800001	12910000	Percent Similarity:	39.9%	Conservative:	49
AP008207_129	12900001	13010000	Best Local Similarity:	31.7%	Mismatches:	98
AP008207_130	13000001	13110000	Query Match:	33.1%	Indels:	263
AP008207_131	13100001	13210000	DB:	15	Gaps:	12
AP008207_132	13200001	13310000	US-10-780-703-2 (1-379) x AP008207_386 (1-110000)			
AP008207_133	13300001	13410000	QY	42	GIuAlaAspHisAspHisMetValMetProGlyPheArgPheHisProThrGlu	61
AP008207_134	13400001	13510000	DB	5196	GAGGTAGATGGGCATGAGTACCTGGTGGATGCGCGGTTCAGGTTCACCCGACGGAG	5255
AP008207_135	13500001	13610000	QY	62	GIuGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGlu	81
AP008207_136	13600001	13710000	DB	5256	GAGGAGCTGATCGAGTTCTACCTCCGCGGAGGTGGAGGCAAGCGCTTCAACGTCGAG	5315
AP008207_137	13700001	13810000	QY	82	LeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuPro	97
AP008207_138	13800001	13910000	DB	5316	CTCATCACCTTCCTCGACCTTACCGCTACGACCCCTGGGAGCT-CCCCGGTAATTAGTT	5374
AP008207_139	13900001	14010000	QY	97	-----	97
AP008207_140	14000001	14110000	DB	5375	AGTTACTCTCCGCTCGATCATATCATGATTGATTAGGAATTGAGCATCTAATAAC	5434
AP008207_141	14100001	14210000	QY	98	-----AlaMetAlaAlaIleGlyGluLysGluTyrPheTyrVal	111
AP008207_142	14200001	14310000	DB	5435	AAGCTCGATGGGGTGCAGCAATGGCGGATAGGGGAGAGAGAGTGTTCCTACGTG	5494
AP008207_143	14300001	14410000	QY	112	ProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyr	131
AP008207_144	14400001	14510000	DB	5495	CCTCGGGACAGGAAGTACAGGAACGGGACCGGCCGACCGGTGACGGGTGCGGGTAC	5554
AP008207_145	14500001	14610000	QY	132	TyrLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeu	151
AP008207_146	14600001	14710000	DB	5555	TGAAGGCGACGGGGCGCCGCGATGATCCGAGCCGAGAACACCGCCCATCGGGCTC	5614
AP008207_147	14700001	14810000	QY	152	LysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrPile	171
AP008207_148	14800001	14910000	DB	5615	AGAAGACGCTTGTCTTCTTACTCCGCGAAGGCCCCCAAGGGCGTCCGACAGCTGATC	5674
AP008207_149	14900001	15010000	QY	172	MetAsnGluTyrArgLeuProHisGluThrGluLysTyrGlnLys	187
AP008207_150	15000001	15110000	DB	5675	ATGAACGAGTATCGCTTCCCGCGCGACACCGCGCTACCAACAGGTACCCATCCAT	5734
AP008207_151	15100001	15210000	QY	187	-----	187
AP008207_152	15200001	15310000	DB	5735	CCATGAATCTACTCGCGCTTCGTTCTTCGATCGATCTCTCTCTCATCCAATTAAT	5794
AP008207_153	15300001	15410000	QY	187	-----	187
AP008207_154	15400001	15510000	DB	5795	TGTAGGGGATTGAATTCTCAGTACTAGGTAGAGGATCTTGTGTGTGTGATTCTGCT	5854
AP008207_155	15500001	15610000	QY	187	-----	187
AP008207_156	15600001	15710000	DB	5855	TGCTTTTAGTTCACCTGCTAGTGGCGCTCATGGTTGCCCATGAATTTAGCGGTAGTA	5914
AP008207_157	15700001	15810000	QY	187	-----	187
AP008207_158	15800001	15910000	DB	5915	TTTAAGAGTGGGTATGATTAGGGTTTTCATGAATTTCTTTTGGGTGCTGATTTTTTT	5974
AP008207_159	15900001	16010000	QY	187	-----	187
AP008207_160	16000001	16110000	DB	5975	TGGGGTGTTCGTCATGAGTATGCATTAATTGAGTGATAGATATGCACCTACTCGACCT	6034
AP008207_161	16100001	16210000	QY	187	-----	187
AP008207_162	16200001	16310000	DB	6035	TGCGATCCCTAGATGCTATATCTTGAATATATATGGGGCGCTAGAAATCCTAGATTCTCAA	6094
AP008207_163	16300001	16410000	QY	187	-----	187
AP008207_164	16400001	16510000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_165	16500001	16610000	QY	187	-----	187
AP008207_166	16600001	16710000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_167	16700001	16810000	QY	187	-----	187
AP008207_168	16800001	16910000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_169	16900001	17010000	QY	187	-----	187
AP008207_170	17000001	17110000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_171	17100001	17210000	QY	187	-----	187
AP008207_172	17200001	17310000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_173	17300001	17410000	QY	187	-----	187
AP008207_174	17400001	17510000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_175	17500001	17610000	QY	187	-----	187
AP008207_176	17600001	17710000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_177	17700001	17810000	QY	187	-----	187
AP008207_178	17800001	17910000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_179	17900001	18010000	QY	187	-----	187
AP008207_180	18000001	18110000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_181	18100001	18210000	QY	187	-----	187
AP008207_182	18200001	18310000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_183	18300001	18410000	QY	187	-----	187
AP008207_184	18400001	18510000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_185	18500001	18610000	QY	187	-----	187
AP008207_186	18600001	18710000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_187	18700001	18810000	QY	187	-----	187
AP008207_188	18800001	18910000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_189	18900001	19010000	QY	187	-----	187
AP008207_190	19000001	19110000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_191	19100001	19210000	QY	187	-----	187
AP008207_192	19200001	19310000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_193	19300001	19410000	QY	187	-----	187
AP008207_194	19400001	19510000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154
AP008207_195	19500001	19610000	QY	187	-----	187
AP008207_196	19600001	19710000	DB	6095	TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAAGATGCTAAGATCTGAAC	6154

```
Qy 188 -----Al 188
Db 6155 TGAATTTTCAGTGAATAAACCCTATCTCTCGCGTCCATGAATTCAGAC 6214
Qy 188 aGluLeSerLeuCysArgValTyrIysArgProGlyValGluAsp---HisProSerVa 207
Db 6215 TGAATATCTCACTATCGCGTGTACAAAGCGCACCGCATCGACAGCGGCATGGCCAGGT 6274
Qy 207 l-ProArg-----SerLeuSerThrArgHisAsnHisAsnS 220
Db 6275 ATCCACGGCGCGGTCTCGCGGCACCTCCCGCGCGCGCGGCACCGGTACAGGACAA 6334
Qy 220 erSerThrSerSerArgLeuAlaLeuArgGln-----GlnGlnHisHisSerSers 237
Db 6335 CAAGCAAGGCTCTCTCGTACGTCAGCGCGCGCGCGCTCCCACTCCGTCCAAGTCCA 6394
Qy 237 erSerAsnHisSerAspAsnLeuAsnAsnAsnAsnAsnAsnLeuAsnLeuGluLysL 257
Db 6395 CCTCCTCAGACGCGAGTGCACGTCGCGCGCGCGCGCGCATCGTACGAGCACCGCGCATG 6454
Qy 257 euSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThr 277
Db 6455 GCGGCACAGGACCGTCCCGCGCGC-----ACCATCAGCAGCAGCAGCAGCTCCACGC 6508
Qy 277 spValThrIleAlaLeuAlaAsnGlnAsnIleTyrArg-----ProMetProTyrAspT 295
Db 6509 TGCCAAAGCCTTGGCGCGGTACTCTGCAGATTCCTCGATGGTTCGGTTCGCGCGCGCA 6568
Qy 295 hrSerAsnAsnThrLeuIleValSerThrArg----- 305
Db 6569 TCAGCAGCAGCAATTTTCAGACAGACTTCGCGCGCGGATTGTACAGCAGTACTCCAGAA 6628
Qy 306 -----AsnHisGlnA 309
Db 6629 CACGAGCGGTGCTTCGCGCTCCACGTACTCGCTGTCACTCAACGCGGCATCCAT 6688
Qy 309 spAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyr----- 325
Db 6689 G-GGCAGCTCCGCGCTGCCATTGATGAGCTCGTGTGGGCCACGCGCAGCCGT 6747
Qy 326 --GlnIleSerAspGlyAlaThr-----LeuMetProGlnT 338
Db 6748 CCTACATTAAACCCCGCGCGCGCAGCACCACTATAGTCAATTCTTCACCTGCTACGA 6807
Qy 338 hrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGly----- 350
Db 6808 CGCGCTCGTGCATCAACCAACGCGCGCGCTCGGAGCAGCAGCGCGCGCGCGG 6867
Qy 351 --ThrIlePro-----AsnAsnAlaLeuTyrAspMetTyrAsnProI 364
Db 6868 CGAGCTGCGGATGTGCTCGCGCGCTTCTCCGACAGGATCTGGGACTGGATATATCCA 6927
Qy 364 leValProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
Db 6928 TCCCGAGCGCGGAGGTAGGATTACAGCACCAGTACTGGATTCAAG 6974
```

Search completed: April 6, 2006, 11:43:26
Job time : 6538 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2006, 09:13:25 ; Search time 12942 Seconds
(without alignments)
11445.992 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctctcc.....ctaattctctttaagtaa 2606

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pt.*

9: gb_ro.*

10: gb_scs.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_btg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2606	100.0	74903	15	AC005312 Arabidops
2	476.2	18.3	534	10	CR381764 Arabidops
3	474.2	18.2	541	10	CR381364 Arabidops
4	467.2	17.9	1062	15	AF360308 Arabidops
5	467.2	17.9	1171	15	BT000874 Arabidops
6	463	17.8	524	10	CR381358 Arabidops
7	456.2	17.5	505	10	CR381827 Arabidops
8	447.8	17.2	530	10	CR381634 Arabidops
9	234	9.0	80577	14	AP007649 Arabidops
10	234	9.0	91718	14	AP007409 Arabidops
11	167.6	6.4	110000	15	AP008207_385
12	167.6	6.4	110000	15	AP008207_386
13	167.6	6.4	150064	15	AP003346 Arabidops
14	167.6	6.4	191022	15	AP003431 Arabidops
15	164	6.3	1212	6	AX654050 Sequence
16	158.6	6.1	1631	15	AK102902 Oryza sat
17	151.8	5.8	110000	15	AP008214 007
18	151.8	5.8	185545	15	AP005657 Oryza sat

19	149.2	5.7	1640	15	AK072275 Oryza sat
20	149.2	5.7	1891	15	AK071052 Oryza sat
21	149.2	5.7	2424	15	AK068393 Oryza sat
22	149	5.7	1250	6	AX654570 Sequence
c 23	135.8	5.2	110000	15	AP008211_203
c 24	135.8	5.2	173074	15	AC137623 Oryza sat
25	132.6	5.1	549	6	AX653764 Sequence
26	126.2	4.8	1406	15	AY672069 Hordeum v
27	120.8	4.6	202169	14	AC149549 Gorilla g
28	118.6	4.6	337	10	AL808204 Arabidops
29	118	4.5	1689	15	AY974352 Glycine m
30	117.2	4.5	1688	15	AY974351 Glycine m
31	117.2	4.5	82893	15	AB022218 Arabidops
c 32	117.2	4.5	82893	15	AC024081 Arabidops
33	112.8	4.3	116862	14	AC160836 Arabidops
34	112.2	4.3	945	6	AX506848 Sequence
35	112.2	4.3	945	6	AX651806 Sequence
c 36	111.8	4.3	82381	15	ATF27G19 Arabidops
c 37	111.8	4.3	200001	15	ATCHRIV67 Arabidops
38	111.6	4.3	1106	15	DQ028771 Glycine m
39	110.4	4.2	1095	15	AB049071 Arabidops
40	110.4	4.2	1126	15	BT005044 Arabidops
41	110.4	4.2	1319	6	CS137982 Sequence
42	110.4	4.2	1719	15	BT004079 Arabidops
43	110.2	4.2	1152	15	DQ028772 Glycine m
c 44	109.2	4.2	44499	15	AB016876 Arabidops
c 45	109	4.2	104204	15	ATT93C5 Arabidops

ALIGNMENTS

RESULT 1

AC005312

LOCUS

DEFINITION

AC005312

AC005312

AC005312.3

HTG.

KEYWORDS

ORGANISM

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

AC005312 Arabidopsis thaliana chromosome 2 clone T16F16 map C1C11A04, complete sequence.

AC005312 GI:20197334

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

Arabidopsis thaliana

REFERENCE 1 (bases 1 to 74903)
AUTHORS Rounsley, S.D., Lin, X., Kaul, S., Shea, T.P., Fujii, C.Y., Mason, T.M., Shen, M., Rinning, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 74903)
AUTHORS Lin, X.
JOURNAL Direct Submission
TITLE Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA

REFERENCE 3 (bases 1 to 74903)
AUTHORS Town, C.D. and Kaul, S.
JOURNAL Direct Submission
TITLE Submitted (27-FEB-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org

COMMENT On Apr 18, 2002 this sequence version replaced gi:6598458.
FEATURES Location/Qualifiers
1..74903
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C11A04"
/clone="T16F16"
/ecotype="Columbia"
complement (<1..2731)
/gene="At2g02220"
/notes="synonym: T16F16.1"
complement (<1..>2731)

```

CDS
/genes="At2g02220"
complement (<1..2731)
/genes="At2g02220"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC78507.3"
/db_xref="GI:20197335"
/translation="MRVRCVIVIFLTELCCFFYSESQTSRCHPHDLALRDLPIA
HLSPKPDGWNSSSDCCNWTGITNSNNTGVRILEGNKKLSGLSGLKDEI
RVNLNRNFITSDIPISFNLKNLOTLIDJSSNDLSGGIPTSINLPALQFDSNNKFN
GSLPSHICHNSTQIRVVKLAVNYFAGFTSGFGKCVLLHLCLGMDLTCNIPEDLFH
LKNLNLGIGENLSGLSRLNLSLRLDVSWNLFSGEIPDVDFELPQLKFLGQ
TKGFIGIPKLSANPSLNLNNSLSGLMLNCTAMALNSLDLGTNRNFRGLPE
NLPDCKRLKNVNLARTFHQVPEKPNFESLSYFSLSNSSLANISSALGILQHCNKL
TLVLTLNPHGEALPDDSLHFEKLVLVVAVNCRLTGSMRNLSSNELQLDLNWR
LTGATFSGIDKALFDLUSNNSFTGEIPKSLTKLESRTSRNISNWFSPDPFPMK
RNESARALQYNOIFGPPTTIELGNHNSGPIWEEFNGNLKHLVFDLKNALSGIPSS
LSGWTLEALDLNRLSGIPVLSLOQLSFLSFVYNNLSGIVPSGGQOTFPNSS
FESNHLCEGHRFPFCSEGTESALIKRRRRSGGDI GMAIGAFSGVFLTLTSLIVLRA
RRSGVDPEIESESNNRKEIGEIKLVLFQSNDKELSYDLDLSTNSPDQANI
CGCGFVYKATLPDGKVAIKKLSGDCGQIEREFAEVETLSRAQHNLYLLRGCF
YKNDRLIYSIMENGSLDYWLHNRDGPALLKWKRLRIAQGAAGLLYLHEGCDPHI
LHRDIKSNILLDENFNSHLADFGLARLMSPYETHVSTDLVGTGLYIPPEYQQA"
complement (1..1865)
/genes="At2g02220"
/notes="Overlap with BAC clone F504 (AC005936:1..1865)."
3226..3247
/rpt_family="AT_rich"
complement (3504..3524)
/rpt_family="AT_rich"
complement (4101..4133)
/rpt_family="AT_rich"
complement (6000..7202)
/genes="At2g02230"
/notes="synonym: T16F16.2"
complement (join(<6000..6237,6295..6555,6649..6767,
6867..7202))
/genes="At2g02230"
complement (join(6000..6237,6295..6555,6649..6767,
6867..7202))
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78508.1"
/db_xref="GI:3894158"
/translation="MSQIHGDSNGSGGGGSRNDEISVTRASRFDALPEDCISKVI
SHTSPRDACVVASVSKSVSAQSDLVWEMFLPSEYSSLVQSANHLSKKEIPLSLAD
NSVLVFGKKSFWVERKASGKCYMLSGMTIIGWDSPAYKMTITPESKFEKVAELR
NVCWFVRGKISGMLSKGTHYSVVVFTKANGRSYGFDLVPVEAGVGFVKVATKKS
VYFESGNASRSATSHYSGISSEEEVEGERGMNVGPKERVDSGSEVELGKFYIN
NGCGDGSDEIISIMETONGNWKSLIIQGIETPERSN"
complement (7966..9209)
/genes="At2g02240"
/notes="synonym: T16F16.3; pseudogene, phloem-specific
lectin"
/pseudo
complement (10262..11417)
/genes="At2g02250"
/notes="synonym: T16F16.4"
complement (join(<10262..10691,10775..10896,11052..>11417))
/genes="At2g02250"
complement (join(10262..10691,10775..10896,11052..11417))
/codon_start=1
/product="lectin-like protein"
/protein_id="AAC78509.1"
/db_xref="GI:3894159"
/translation="MGHKQSDVRGKRGKVPKVGSSSMVQKRVHSSGGAIIISGSLFDN
LPDSDISNIIISFSPRDACVAASVKTFESAVNSDSVDKFLPESDSSLVPPSRVFS
KRELFAICDNPLVLDGGKSWLEKENGKCFMLSPKSKMTITWCTFYWRWISIP
EAEFEEVPELNNVCFEVRGGMTKELSGTGRYSAYIVFKTKGCSNGLGDPVEATVG
LVQSESSORHI YFVGPSDQRDRDTRDVTPTKRRKDGWMEALGQFFNESGCDVDTS
ILBKIPTYKRLIIQGIETPYSFYI"

```

```

gene
complement (12061..13367)
/genes="At2g02260"
/notes="synonym: T16F16.5; pseudogene, reverse
transcriptase"
/pseudo
complement (13604..15132)
/genes="At2g02270"
/notes="synonym: T16F16.6"
complement (join(<13604..13663,13728..14100,14227..14312,
14854..>15132))
/genes="At2g02270"
complement (join(13604..13663,13728..14100,14227..14312,
14854..15132))
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78510.1"
/db_xref="GI:3894160"
/translation="MAQGEIVPRSPFDGLPENCISNIISFTTPRDACFAASVSKAFE
SAVQSDSVMEKFLPDYSSIVPESRVFLSKELCFSLCRVPLLIIEGKKSFWDKTSG
BKCIWLSKGMVIVSNVSPQEPVEVPELLYDSHFEVCGRLSTKYLSPRTSYVIVFT
NLYFGVLTLEPPFRVVRVPGFTDLKVEREYVTRPKREKKNWAEALCEFFNETSCDV
EVSVDNSYWKSGLVIQGIETPRTKKHQVKKKNCIVNILRFFYCNV"
complement (16098..17046)
/genes="At2g02280"
/notes="synonym: T16F16.7"
complement (join(<16098..16334,16849..>17046))
/genes="At2g02280"
complement (join(16098..16334,16849..17046))
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78511.1"
/db_xref="GI:3894161"
/translation="WNTQLSOKTRYSAIVVTKTYRPHGFKHIGVGTGHGTPKAKR
WERKLDGNDLWLGCKKFKAKQKQFYSNKSTFDKPIHLIKLEEGEDGMATEFGF
FAEGGGLDDCDEIVLSVIDYVAYWKGLIIQGIIDIRPTKSP"
16391..16419
/rpt_family="AT_rich"
16523..16558
/rpt_family="AT_rich"
complement (16628..16649)
/rpt_family="(TA)n"
16672..16711
/rpt_family="(A)n"
complement (16770..16793)
/rpt_family="AT_rich"
complement (17280..17307)
/rpt_family="AT_rich"
complement (17616..17643)
/rpt_family="AT_rich"
complement (17834..17896)
/rpt_family="AT_rich"
complement (18077..18151)
/rpt_family="(GAA)n"
complement (18618..18659)
/rpt_family="AT_rich"
19173..19192
/rpt_family="(A)n"
20060..20968
/genes="At2g02290"
/notes="synonym: T16F16.8; predicted by genscan and
genefinder"
<20060..>20968
/genes="At2g02290"
20060..20968
/genes="At2g02290"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC78512.1"
/db_xref="GI:3894162"
/translation="MSLVENLSLEPQPKRKIDPEYKESPSRNCSHWFVRNKCASCN
TIVDNYQGRSPFDYLTGTIGMSNEALGFTKRLISQTSWLEDKHLVLDLDTLVHTIK

```

VSQSESEKYYITEEVESRKDLRRFNTGFPPEESLIKLRSFVHQFLKECNEMFSLVYVTK
GGYDAQVLEIMIDPDKIYGNRVITRRSPGFKTLDLVLADRGIVVDDKSSWPH
DKNLLQIARYKYFDQSGCLLSECKKIDESDEKPLNTALRFLMDVHEEFPDWSRK
DLDSDVDRPLAKEISLRWKRKADGVV"
complement (22168..23432)
/gene="At2902300"
/note="synonym: T16F16.9"
complement (join (<22168..22558,22680..22798,23088..>23432))
/gene="At2902300"
complement (join (22168..22558,22680..22798,23088..23432))
/gene="At2902300"
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78513.1"
/db_xref="GI:3894163"

Query Match 100.0%; Score 2606; DB 15; Length 74903;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCAATTGTATCTCCACAAACAGCATCATTTCCCATGAGTAACCAAGTCAACAATAAC	60
DB	63651	ATGGCAATTGTATCTCCACAAACAGCATCATTTCCCATGAGTAACCAAGTCAACAATAAC	63710
QY	61	GAAGAAAGGTATAGAGACAAATGATCATAGAGCGGCGCAAGAGAGTCATGTCACAAATGAA	120
DB	63711	GAAGAAAGGTATAGAGACAAATGATCATAGAGCGGCGCAAGAGAGTCATGTCACAAATGAA	63770
QY	121	GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT	180
DB	63771	GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT	63830
QY	181	GAAGAAAGCTCATAGAGTTTACCTTCGCGAAAGTTGAAGCAAGCGCTTAAATGTA	240
DB	63831	GAAGAAAGCTCATAGAGTTTACCTTCGCGAAAGTTGAAGCAAGCGCTTAAATGTA	63890
QY	241	GAATCATACATCT	300
DB	63891	GAATCATACATCT	63950
QY	301	ACATTCACATAACACACATAAATCATCTCAAACTATTTGGAAATCTTAAATTTCTATTC	360
DB	63951	ACATTCACATAACACACATAAATCATCTCAAACTATTTGGAAATCTTAAATTTCTATTC	64010
QY	361	TATGTTAAGATCTTTCT	420
DB	64011	TATGTTAAGATCTTTCT	64070
QY	421	ATATATGTACCTACCT	480
DB	64071	ATATATGTACCTACCT	64130
QY	481	CTTTAAAGCGTCT	540
DB	64131	CTTTAAAGCGTCT	64190
QY	541	TATATATCATAGATATACAAATATGTTATGTTTTCATATATGTTATGTTATGTTATGTT	600
DB	64191	TATATATCATAGATATACAAATATGTTATGTTTTCATATATGTTATGTTATGTTATGTT	64250
QY	601	TCATAGATATATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTT	660
DB	64251	TCATAGATATATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTT	64310
QY	661	GTGCCAAGAGATCGGAATATAGAAATGGAGATAGACCGGAAACCGAGTAAACGACTTCAG	720
DB	64311	GTGCCAAGAGATCGGAATATAGAAATGGAGATAGACCGGAAACCGAGTAAACGACTTCAG	64370
QY	721	TATTTGGAAGCCACCGAGCTGATAGATGATGATGATGATGATGATGATGATGATGATGAT	780
DB	64371	TATTTGGAAGCCACCGAGCTGATAGATGATGATGATGATGATGATGATGATGATGATGAT	64430
QY	781	TTAAAGAAACCCCTAGTTTCT	840

DB	64431	TTAAAGAAACCCCTAGTTTCT	64490
QY	841	ATCATGAACGAGTATCGTCTTCGCGACCATGAAACCGGAGAGTACCAAAAGGTATAAATTT	900
DB	64491	ATCATGAACGAGTATCGTCTTCGCGACCATGAAACCGGAGAGTACCAAAAGGTATAAATTT	64550
QY	901	CTACTATAAATCTCTATATATATCTTATTCATACATACATAGATATACCCCTAGCTAGTGG	960
DB	64551	CTACTATAAATCTCTATATATATCTTATTCATACATACATAGATATACCCCTAGCTAGTGG	64610
QY	961	TGAGGCGCTTTAAAAATGAAATTAATCCCTAGACAGTTTGAATTTTCTTTTCTGACTAG	1020
DB	64611	TGAGGCGCTTTAAAAATGAAATTAATCCCTAGACAGTTTGAATTTTCTTTTCTGACTAG	64670
QY	1021	TTTTTATTTATTTTGGAAATTTGAAATTTGATTCGATAAGATCAAAAATACCTTGTGAATGA	1080
DB	64671	TTTTTATTTATTTTGGAAATTTGAAATTTGATTCGATAAGATCAAAAATACCTTGTGAATGA	64730
QY	1081	ATGTCAGGCGGCTTTGGCGCTTAAATCCAGAAAAATGTTTCATGTCATATCGGTGAATCT	1140
DB	64731	ATGTCAGGCGGCTTTGGCGCTTAAATCCAGAAAAATGTTTCATGTCATATCGGTGAATCT	64790
QY	1141	TTAAATTTGCTAGACATGGCCCATATGTTTATAGTAGAATACATTAATAGATAGATGCATAC	1200
DB	64791	TTAAATTTGCTAGACATGGCCCATATGTTTATAGTAGAATACATTAATAGATAGATGCATAC	64850
QY	1201	ACATATATATAACACACACAGTATCACACTCGACATTCATATACCTTAAATTTCTGCAGAGA	1260
DB	64851	ACATATATATAACACACACAGTATCACACTCGACATTCATATACCTTAAATTTCTGCAGAGA	64910
QY	1261	CATAGTTAGTTTCTTTTACAAATTTATGACATGAATGTTCTCTCTCTCTCTCTCTCTCTCT	1320
DB	64911	CATAGTTAGTTTCTTTTACAAATTTATGACATGAATGTTCTCTCTCTCTCTCTCTCTCTCT	64970
QY	1321	CATGCTCTTCTATTTAAGTTTACCCCAACATTTTGTGAAATTAATTTGGCATATATGAATATA	1380
DB	64971	CATGCTCTTCTATTTAAGTTTACCCCAACATTTTGTGAAATTAATTTGGCATATATGAATATA	65030
QY	1381	CCAAACATATTTATGCGAACATTTAAATCTATACGAATGATACGGTTTATGGAGTAG	1440
DB	65031	CCAAACATATTTATGCGAACATTTAAATCTATACGAATGATACGGTTTATGGAGTAG	65090
QY	1441	ACCGAAAAATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1500
DB	65091	ACCGAAAAATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	65150
QY	1501	TCCTTTCGACTTATATGTCGTCACCATTTGAAACCATATAATTTTATAAAATTTTCTATGAT	1560
DB	65151	TCCTTTCGACTTATATGTCGTCACCATTTGAAACCATATAATTTTATAAAATTTTCTATGAT	65210
QY	1561	ATATATGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1620
DB	65211	ATATATGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	65270
QY	1621	TATGTCATGTCATGTCACATTCCTGTCATCTATATAAACTAAATTAACGCAATTAATAT	1680
DB	65271	TATGTCATGTCATGTCACATTCCTGTCATCTATATAAACTAAATTAACGCAATTAATAT	65330
QY	1681	ATAAAAAATATAAATAATATCATCTTAAATTTATTTTGGCATTTGTAGCTCATATGATAGT	1740
DB	65331	ATAAAAAATATAAATAATATCATCTTAAATTTATTTTGGCATTTGTAGCTCATATGATAGT	65390
QY	1741	ACTCTAAATTTCTTTTAAACGTCGCTATCTTTTGTCTAAATGCTTAACTTTTACATAGTTTG	1800
DB	65391	ACTCTAAATTTCTTTTAAACGTCGCTATCTTTTGTCTAAATGCTTAACTTTTACATAGTTTG	65450
QY	1801	TGAATCTTCTTTCAAAACCATATCTTCGATAAATGATATTTTTCATAGATATCTTAGTC	1860
DB	65451	TGAATCTTCTTTCAAAACCATATCTTCGATAAATGATATTTTTCATAGATATCTTAGTC	65510
QY	1861	TATATTTGATAAATTTTGATATATGATCAAGTCTCTAATCAATGTCTCATGTATAAATAT	1920

AUTHORS Clarke, J.H.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAR-2004) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK

COMMENT AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
 On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock code: N165877.

FEATURES
 Location/Qualifiers
 source 1..505
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /clone="AC005312"
 /ecotype="Landsberg erecta NASC stock code NW20"
 /notes="Unknown"
 STS 1..505
 /standard_name="GT_5.102321"

ORIGIN
 Query Match 17.5%; Score 456.2; DB 10; Length 505;
 Best Local Similarity 97.6%; Pred. No. 3.3e-65;
 Matches 495; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

964 GGCCTTTAAATGGAATTAATCCCTAGACAGTTGGAATTTTTC-TTTTTGACTAGTT 1022
 Db 1 GGCCTTTAAATGGAATTAATCCCTAGACAGTTGGAATTTTCTTTTTCGACTAGTT 60

1023 TTATTTATTTATTTGGAATTTGATTCGATAAGATCAAAAATATCTGTGAATGCACTAAAT 1082
 Db 61 TTATTTATTTATTTGGAATTTGATTCGATAAGATCAAAAATATCTGTGAATGCACTAAAT 120

1083 GTCAGCGGGGTTTGGCTTTAAATCCAG-NAAATGTTTCATGTCATATGGTGCACTTT 1141
 Db 121 GTCAGCGGGGTTTGGCTTTAAATCCAGAAAAAATGTTTCATGTCATATGGTGCACTTT 180

1142 TAAATGCTAGACATGCCCATATGTTATAGTAGATACATAATTAATAGATAGATGATACA 1201
 Db 181 TAAATGCTAGACATGCCCATATGTTATAGTAGATACATAATTAATAGATAGATGATACA 240

1202 CATATATATAACACACAAGATATCACCTGCACATTCATATACCTTTAAATTTCTGCAGAG 1261
 Db 241 C--ATATATAACACACAAGATATCACCTGCACATTCATATACCTTTAAATTTCTGCAGAG 298

1262 ATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCTGCTCTTCTCTCACATTAATTC 1321
 Db 299 ATAGTTAGTTTCTTACAAATTTATTCATGATGAATGTTCTGCTCTTCTCTCACATTAATTC 358

1322 ATGCTTCTATTTAAAGTTACCAACATTTTGGAAATAATTTGGCATATATGAATTTATAC 1381
 Db 359 ATGCTTCTATTTAAAGTTACCAACATTTTGGAAATAATTTGGCATATATGAATTTATAC 418

1382 CAACATATTTATATGCGAACATTTTAAATCTTATACGAATGATACCGTTTTATGGAGTAGA 1441
 Db 419 CAACATATCTATATGCGAACATTTTAAATCTTATACGAAGATTAACCGTTTTATGGAGTAGA 478

1442 CCGAAAAAATATTTATGATATACGGAAAA 1468
 Db 479 CCGAAAAAATATCTATGATATACGGATAA 505

RESULT 8
 CR381634
 LOCUS Arabidopsis thaliana transposon insertion STS GT_5.101721, sequence tagged site.
 DEFINITION CR381634
 ACCESSION CR381634
 VERSION CR381634.1 GI:45773115

KEYWORDS STS, sequence tagged site.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 AUTHORS Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Walsh, S., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 530)
 AUTHORS Clarke, J.H.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAR-2004) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK

COMMENT AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
 On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock code: N165391.

FEATURES
 Location/Qualifiers
 source 1..530
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /clone="AC005312"
 /ecotype="Landsberg erecta NASC stock code NW20"
 /note="Unknown"
 STS 1..530
 /standard_name="GT_5.101721"

ORIGIN
 Query Match 17.2%; Score 447.8; DB 10; Length 530;
 Best Local Similarity 96.8%; Pred. No. 7.8e-64;
 Matches 489; Conservative 0; Mismatches 12; Indels 4; Gaps 3;

956 GGTGTGTAGGGCTTTAAATTTGAAATTAATCCCTAGACAGTTTGAATTTTTC-TTTTTT 1014
 Db 28 GGTGTGTAGGGCTTTAAATTTGAAATTAATCCCTAGACAGTTTGAATTTTTC-TTTTTT 87

1015 GACTAGTTTATTTATTTATTTTGGAAATGATTCGATAGATCAAAAATACCTGTGAATG 1074
 Db 88 GACTAGTTTATTTATTTATTTTGGAAATGATTCGATAGATCAAAAATACCTGTGAATG 147

1075 GACTAAATGTCAGCGGGCTTTGCGCTTAAATCCA-GAAAAATGTTTCATGTCATATCGCT 1133
 Db 148 GACTAAATGTCAGCGGGCTTTGCGCTTAAATCCA-TAAAAATGTTTCATGTCATATCGCT 207

1134 GAACCTCTTTAAATTCGTAGACATGCGCCCATATGTTATAGTAGAATACATTAATAGATAGA 1193
 Db 208 GAACCTCTTTAAATTCGTAGACATGCGCCCATATGTTATAGTAGAATACATTAATAGATAGA 267

1194 TGCATACACATATATATAACACACAGATATCACCTCGACATTCATATACCTTAATTCCT 1253
 Db 268 TGCATACAC--ATATATAAACACACAGATATCACCTCGACATTCATATACCTTAATTCCT 325

1254 GCAGAGACATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCTGCTCTTCTCTCAC 1313
 Db 326 GCAGAGACATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCTGCTCTTCTCTCAC 385

1314 ATTAATTCATGTTCTTCTATTTAAAGTTACCCACATTTTGAATTTTGGCATATATG 1373
 Db 386 ATTAATTCATGTTCTTCTATTTAAAGTTACCCACATTTTGAATTTTGGCATATATG 445

1374 AATTATACCAACATATTTATATGCGACATTTAAATCTATACGATGATTAACGGTTTAT 1433
 Db 446 AATTATACCAACATATCTATATGCGAACACTTAAAAATCTATACGAGGATTAACGGTTTAT 505

1434 GGAGTAGACCGGAAAAAATATTTATGT 1458

AP008207_063 6300001 6410000
 AP008207_064 6400001 6510000
 AP008207_065 6500001 6610000
 AP008207_066 6600001 6710000
 AP008207_067 6700001 6810000
 AP008207_068 6800001 6910000
 AP008207_069 6900001 7010000
 AP008207_070 7000001 7110000
 AP008207_071 7100001 7210000
 AP008207_072 7200001 7310000
 AP008207_073 7300001 7410000
 AP008207_074 7400001 7510000
 AP008207_075 7500001 7610000
 AP008207_076 7600001 7710000
 AP008207_077 7700001 7810000
 AP008207_078 7800001 7910000
 AP008207_079 7900001 8010000
 AP008207_080 8000001 8110000
 AP008207_081 8100001 8210000
 AP008207_082 8200001 8310000
 AP008207_083 8300001 8410000
 AP008207_084 8400001 8510000
 AP008207_085 8500001 8610000
 AP008207_086 8600001 8710000
 AP008207_087 8700001 8810000
 AP008207_088 8800001 8910000
 AP008207_089 8900001 9010000
 AP008207_090 9000001 9110000
 AP008207_091 9100001 9210000
 AP008207_092 9200001 9310000
 AP008207_093 9300001 9410000
 AP008207_094 9400001 9510000
 AP008207_095 9500001 9610000
 AP008207_096 9600001 9710000
 AP008207_097 9700001 9810000
 AP008207_098 9800001 9910000
 AP008207_099 9900001 10010000
 AP008207_100 10000001 10110000
 AP008207_101 10100001 10210000
 AP008207_102 10200001 10310000
 AP008207_103 10300001 10410000
 AP008207_104 10400001 10510000
 AP008207_105 10500001 10610000
 AP008207_106 10600001 10710000
 AP008207_107 10700001 10810000
 AP008207_108 10800001 10910000
 AP008207_109 10900001 11010000
 AP008207_110 11000001 11110000
 AP008207_111 11100001 11210000
 AP008207_112 11200001 11310000
 AP008207_113 11300001 11410000
 AP008207_114 11400001 11510000
 AP008207_115 11500001 11610000
 AP008207_116 11600001 11710000
 AP008207_117 11700001 11810000
 AP008207_118 11800001 11910000
 AP008207_119 11900001 12010000
 AP008207_120 12000001 12110000
 AP008207_121 12100001 12210000
 AP008207_122 12200001 12310000
 AP008207_123 12300001 12410000
 AP008207_124 12400001 12510000
 AP008207_125 12500001 12610000
 AP008207_126 12600001 12710000
 AP008207_127 12700001 12810000
 AP008207_128 12800001 12910000
 AP008207_129 12900001 13010000
 AP008207_130 13000001 13110000
 AP008207_131 13100001 13210000
 AP008207_132 13200001 13310000
 AP008207_133 13300001 13410000
 AP008207_134 13400001 13510000
 AP008207_135 13500001 13610000

AP008207_136 13600001 13710000
 AP008207_137 13700001 13810000
 AP008207_138 13800001 13910000
 AP008207_139 13900001 14010000
 AP008207_140 14000001 14110000
 AP008207_141 14100001 14210000
 AP008207_142 14200001 14310000
 AP008207_143 14300001 14410000
 AP008207_144 14400001 14510000
 AP008207_145 14500001 14610000
 AP008207_146 14600001 14710000
 AP008207_147 14700001 14810000
 AP008207_148 14800001 14910000
 AP008207_149 14900001 15010000
 AP008207_150 15000001 15110000
 AP008207_151 15100001 15210000
 AP008207_152 15200001 15310000
 AP008207_153 15300001 15410000
 AP008207_154 15400001 15510000
 AP008207_155 15500001 15610000
 AP008207_156 15600001 15710000
 AP008207_157 15700001 15810000
 AP008207_158 15800001 15910000
 AP008207_159 15900001 16010000
 AP008207_160 16000001 16110000
 AP008207_161 16100001 16210000
 AP008207_162 16200001 16310000
 AP008207_163 16300001 16410000
 AP008207_164 16400001 16510000
 AP008207_165 16500001 16610000
 AP008207_166 16600001 16710000
 AP008207_167 16700001 16810000
 AP008207_168 16800001 16910000
 AP008207_169 16900001 17010000
 AP008207_170 17000001 17110000
 AP008207_171 17100001 17210000
 AP008207_172 17200001 17310000
 AP008207_173 17300001 17410000
 AP008207_174 17400001 17510000
 AP008207_175 17500001 17610000
 AP008207_176 17600001 17710000
 AP008207_177 17700001 17810000
 AP008207_178 17800001 17910000
 AP008207_179 17900001 18010000
 AP008207_180 18000001 18110000
 AP008207_181 18100001 18210000
 AP008207_182 18200001 18310000
 AP008207_183 18300001 18410000
 AP008207_184 18400001 18510000
 AP008207_185 18500001 18610000
 AP008207_186 18600001 18710000
 AP008207_187 18700001 18810000
 AP008207_188 18800001 18910000
 AP008207_189 18900001 19010000
 AP008207_190 19000001 19110000
 AP008207_191 19100001 19210000
 AP008207_192 19200001 19310000
 AP008207_193 19300001 19410000
 AP008207_194 19400001 19510000
 AP008207_195 19500001 19610000
 AP008207_196 19600001 19710000

Query Match 6.4%; Score 167.6; DB 15; Length 110000;
 Best Local Similarity 71.3%; Pred. No. 3.3e-18;
 Matches 221; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 Qy 605 AGATATATGTATATGCGGCGGATAGGAGAGAGAGTGTCTTCTATGTGC 664
 Db 105436 AGTCGATGGGGTGCAGCAATGCGGCGGATAGGAGAGAGTGTCTTCTATGTGC 105495
 Qy 665 CAAGAGATCGGAATATAGAAATGGAGATAGACCGAACCGAGTACAGTATTT 724
 Db 105496 CTCGGGACAGAGTACAGACCGGCGGCGGCGGCGGCGGCGGCGGCTACT 105555

Qy 725 GGAAGCCACCGAGCTGATAGGATGATCGAGACTTCTCGGCTATCGGATTAA 784
|||||
Db 105556 GGAAGCCACCGGGCCGACCGGATGATCCGAGCCGAGACACCGCCCATCGGCTCA 105615
Qy 785 AGAAACCTAGTTTCTACTCTGTGTAAGCCCTAAAGGCACTCGTACTAGTTGGATCA 844
|||||
Db 105616 AGAAGACCTGTCTTCTACTCCGCAAGGGCCCAAGGGGTCGCGACGAGCTGGATCA 105675
Qy 845 TGAACGAGTATCGTCTTCCGACCATGAAACCGAGAGTACCAAGAGTATAAATTCCTAC 904
|||||
Db 105676 TGAACGAGTATCGCTTCTCCCGCCGACCGACCGGTACCAAGAGTACCCATCCATC 105735
Qy 905 TATAACTCTA 914
|||||
Db 105736 CATGAATCTA 105745

RESULT 12
AP008207_386
WPCOMMENT

Sequence split into 433 fragments LOCUS AP008207 Accession AP008207

Fragment Name	Begin	End
AP008207_000	1	110000
AP008207_001	100001	210000
AP008207_002	200001	310000
AP008207_003	300001	410000
AP008207_004	400001	510000
AP008207_005	500001	610000
AP008207_006	600001	710000
AP008207_007	700001	810000
AP008207_008	800001	910000
AP008207_009	900001	1010000
AP008207_010	1000001	1110000
AP008207_011	1100001	1210000
AP008207_012	1200001	1310000
AP008207_013	1300001	1410000
AP008207_014	1400001	1510000
AP008207_015	1500001	1610000
AP008207_016	1600001	1710000
AP008207_017	1700001	1810000
AP008207_018	1800001	1910000
AP008207_019	1900001	2010000
AP008207_020	2000001	2110000
AP008207_021	2100001	2210000
AP008207_022	2200001	2310000
AP008207_023	2300001	2410000
AP008207_024	2400001	2510000
AP008207_025	2500001	2610000
AP008207_026	2600001	2710000
AP008207_027	2700001	2810000
AP008207_028	2800001	2910000
AP008207_029	2900001	3010000
AP008207_030	3000001	3110000
AP008207_031	3100001	3210000
AP008207_032	3200001	3310000
AP008207_033	3300001	3410000
AP008207_034	3400001	3510000
AP008207_035	3500001	3610000
AP008207_036	3600001	3710000
AP008207_037	3700001	3810000
AP008207_038	3800001	3910000
AP008207_039	3900001	4010000
AP008207_040	4000001	4110000
AP008207_041	4100001	4210000
AP008207_042	4200001	4310000
AP008207_043	4300001	4410000
AP008207_044	4400001	4510000
AP008207_045	4500001	4610000
AP008207_046	4600001	4710000
AP008207_047	4700001	4810000
AP008207_048	4800001	4910000
AP008207_049	4900001	5010000

AP008207_050 5000001 5110000
AP008207_051 5100001 5210000
AP008207_052 5200001 5310000
AP008207_053 5300001 5410000
AP008207_054 5400001 5510000
AP008207_055 5500001 5610000
AP008207_056 5600001 5710000
AP008207_057 5700001 5810000
AP008207_058 5800001 5910000
AP008207_059 5900001 6010000
AP008207_060 6000001 6110000
AP008207_061 6100001 6210000
AP008207_062 6200001 6310000
AP008207_063 6300001 6410000
AP008207_064 6400001 6510000
AP008207_065 6500001 6610000
AP008207_066 6600001 6710000
AP008207_067 6700001 6810000
AP008207_068 6800001 6910000
AP008207_069 6900001 7010000
AP008207_070 7000001 7110000
AP008207_071 7100001 7210000
AP008207_072 7200001 7310000
AP008207_073 7300001 7410000
AP008207_074 7400001 7510000
AP008207_075 7500001 7610000
AP008207_076 7600001 7710000
AP008207_077 7700001 7810000
AP008207_078 7800001 7910000
AP008207_079 7900001 8010000
AP008207_080 8000001 8110000
AP008207_081 8100001 8210000
AP008207_082 8200001 8310000
AP008207_083 8300001 8410000
AP008207_084 8400001 8510000
AP008207_085 8500001 8610000
AP008207_086 8600001 8710000
AP008207_087 8700001 8810000
AP008207_088 8800001 8910000
AP008207_089 8900001 9010000
AP008207_090 9000001 9110000
AP008207_091 9100001 9210000
AP008207_092 9200001 9310000
AP008207_093 9300001 9410000
AP008207_094 9400001 9510000
AP008207_095 9500001 9610000
AP008207_096 9600001 9710000
AP008207_097 9700001 9810000
AP008207_098 9800001 9910000
AP008207_099 9900001 10010000
AP008207_100 10000001 10110000
AP008207_101 10100001 10210000
AP008207_102 10200001 10310000
AP008207_103 10300001 10410000
AP008207_104 10400001 10510000
AP008207_105 10500001 10610000
AP008207_106 10600001 10710000
AP008207_107 10700001 10810000
AP008207_108 10800001 10910000
AP008207_109 10900001 11010000
AP008207_110 11000001 11110000
AP008207_111 11100001 11210000
AP008207_112 11200001 11310000
AP008207_113 11300001 11410000
AP008207_114 11400001 11510000
AP008207_115 11500001 11610000
AP008207_116 11600001 11710000
AP008207_117 11700001 11810000
AP008207_118 11800001 11910000
AP008207_119 11900001 12010000
AP008207_120 12000001 12110000
AP008207_121 12100001 12210000
AP008207_122 12200001 12310000

	TITLE	The genome sequence and structure of rice chromosome 1 Nature 420 [6913], 312-316 (2002)
JOURNAL	PUBMED	12447438
REFERENCE	AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	JOURNAL	Submitted (21-MAR-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kamonndai 2-1-2, Tsukuba, Ibaraki, 305-8602, Japan (E-mail:tsasaikubias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT		On May 15, 2002 this sequence version replaced gi:13429996. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), GENESK (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://opal.tigr.org/tdb/glimmerm/glrm form.html), RicheHM (http://rgp.dna.affrc.go.jp/RicheHM/) -SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/doc/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blat/db/) and the cDNA regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using Full-length cDNAs represent the identified accession no. and RGP clone ID. BLASTN with the corresponding DBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of BI099D03 clone has an overlap with P0434C04 (DBJ: AP003407) clone at 3' end and an overlap with BI078G07 (DBJ: AP003407) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomesSeq.html .
FEATURES	source	1. 191022 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="1" /clones="BI099D03"
	gene	join(499..854,897..1173,1427..1509,1733..1864,1953..2025,2270..2482) /gene="BI099D03.1" /note="retroelement-like protein"
	misc_feature	join(499..854,897..1173,1427..1509,1733..1864,1953..2025,2270..2482) /note="BIO99D03.1"
	gene	/complement(3633..3800) /gene="BI099D03.2"
	misc_feature	complement(3633..3800) /gene="BIO99D03.2"
	gene	/note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" complement(join(4088..4327,4393..4515,4599..4709,4806..5046,5275..5621)) /gene="BIO99D03.3"
	misc_feature	complement(join(4088..4327,4393..4515,4599..4709,4806..5046,5275..5621))
misc_feature		23069..24124 /genes="P0434C04.7-2" /notes="contains full-length cDNA(s): AKI20493 probably inactive due to 5'UTR too long in CDS" join(23140..23674,24261..24346,25033..25178,25498..25709).25825,26539..26682,26780..27409) /genes="P0434C04.7-1" /notes="supported by full-length cDNA(s): AKI04828" join(23565..23674,24261..24346,25033..25178,25498..25709).25825,26539..26682,26780..26932) /genes="P0434C04.7-1" /notes="contains EST(s): D40498(S25231),AU095566(R4091),AU033072(S2523)
Query Match Best Local Similarity Matches 22;		Conservative 71.3%; Score 167.6; DB 15; Length 150064; Pred. No. 3.le-18; Mismatches 89; Indels 0; Gaps 0;
QY	605	AGATATATGTTATGTATCGTTCGGGGCATAGGAGAAGAGTGCTTCATTGTGC 664
Db	126117	AGCTCGATGGGGTGCCAGCAATCGGCCGATAGGGGGAAGAGTGGTCTTCACGTGC 126176
QY	665	CAGAGATCGGAATATAGAAATGAGTAGACCGAACCGACTTAACGATTCAGGNATT 724
Db	126177	CTCGGACAGGAAGTACAGGAACCGGACCGGCCAACCCTCGGGGTACT 126236
QY	725	GGAAGCACCGGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 784
Db	126237	GGAAGGCGGCGGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 126296
QY	785	AGAAAACCCCTAGTTTTCTACTCTGTAAGCCCCATAAGGCACTCGTACTAGTTGATCA 844
Db	126297	AGAAGAGCGTTGTCTTCTACTCGGCAAGCCCAGGCGTCCGACAGCTGATCA 126356
QY	845	TGAACGATGATGCTTCCGACCATGAAACCGAAGTACCAAAGGTATAAATTCATC 904
Db	126357	TGAACGATGATGCTTCCGACCATGAAACCGAAGTACCAAAGGTATAAATTCATC 126416
QY	905	TATAACTCTA 914
Db	126417	CATGAATCTA 126426
RESULT 14		
AP003431		
LOCUS	AP003431	Oryza sativa (japonica cultivar-group) linear PLN 22-DRC-2004
DEFINITION	BAC clone:BI099D03	
ACCESSION	AP003431	BA000010
VERSION	AP003431.2	GI:20804950
KEYWORDS		Oryza sativa (japonica cultivar-group)
SOURCE		Oryza sativa (japonica cultivar-group)
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartioideae; Oryzeae; Oryza.
REFERENCE		1
AUTHORS		Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T., Katayose,Y., Wu,J., Niimura,Y., Cheng,Z., Nagamura,Y., Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arioka,K., Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K., Hamada,M., Harada,C., Hijiishi,S., Honda,M., Ichikawa,Y., Idonuma,A., Itijima,M., Ikeda,M., Ikono,M., Ito,S., Ito,T., Ito,Y., Ito,Y., Iwabuchi,A., Kamiya,K., Karasawa,W., Katsugiri.S., Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara.T., Mizuno.H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima.M., Nakana.Y., Nakamichi.Y., Nakamura,M., Namiki.N., Negishi.M., Ohta.I., Ono,N., Saiji,S., Sekai.K., Shibata,M., Shimokawa,T., Shomura.A., Song.J., Takazaki.Y., Terasawa,K., Tsuji,K., Waki,K., Yamagata.H., Yamanoe,H., Yoshiki.S., Yoshihara,R., Yukawa.K., Zhong,H., Iwama,H., Endo.T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y., Vano.W., Jiang.J. and Gojobori.T.

JOURNAL Patent: WO 03000898-A 3920 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
source 1. .1212
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

	Query Match	6.3%	Score 164;	DB 6;	Length 1212;
	Best Local Similarity	70.8%;	Pred. No. 3.7e-17;		
	Matches 218;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;
QY	605	AGATATATGATATGACGATATGCGCGGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGC	664		
Db	161	AGCTCGATGGGGTGCAGCAATGGCGCGATAGGGGAGAGGAGTGGTTCTTCTACGTGC	220		
QY	665	CAAGAGATCGGAATATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGGATATT	724		
Db	221	CTCGGACACAGGAAGTACAGAACCGGGACCGGCCGAAACCGGGTGACGGCGTCGGGGTACT	280		
QY	725	GGAAAGCCACCGGAGCTGATAGGATGATCAGATCGGAGACTTCTCGGCCCTATCGGATTAA	784		
Db	281	GGAGCGCGACGGGGCGGACCGGATGATCCGAGCCCGAGAACACCGCCCATCGGGCTCA	340		
QY	785	AGAAACCCCTAGTTTCTACTCTGGTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCA	844		
Db	341	AGAAGACGCTTGTCTTCTACTCCGGCAAGGCCCCCAAGGGCGTCCCGCAGCCTGGATCA	400		
QY	845	TGAACGAGTATCGTCTTCCGACCATGAAACCGAGAAAGTACCAAAAGGTATAAATCTTAC	904		
Db	401	TGAACGAGTATCGCCTTCCCGCCGACACCGACCGCTACCAAGACTGAAATCTCAC	460		
QY	905	TATAACTC	912		
Db	461	TATGCCGC	468		

Search completed: April 6, 2006, 12:58:34
Job time : 12948 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocorelation Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 10:14:38 ; Search time 983 Seconds
(without alignments)
3188.295 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MATVSTTIIIPMSNQNNN.....MNPPIVDGNDHDTNIPFK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlh
-USER=US10780703 -POOL=US10780703/runat_06042006_101348_6580/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss07
-USER=US10780703 @CGN 1.1.1026 @runat_06042006_101348_6580 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	1140	8	US-10-780-703-1
2	2006.5	98.6	1444	6	US-10-225-066A-115
3	2006.5	98.6	1444	7	US-10-225-067-87
4	2006.5	98.6	1444	7	US-10-374-780A-2475
5	2006.5	98.6	1444	9	US-10-225-066A-115
6	1735	85.3	2606	8	US-10-780-703-3
7	889	43.7	785	7	US-10-424-599-121054

8	812	39.9	1433	7	US-10-437-963-88434	Sequence 88434, A
9	781.5	38.4	1291	8	US-10-425-115-10100	Sequence 10100, A
10	781	38.4	743	7	US-10-437-963-81769	Sequence 81769, A
11	737	36.2	792	8	US-10-425-115-173486	Sequence 173486, A
12	545	26.8	1321	7	US-10-437-963-53771	Sequence 100253, A
13	535.5	26.3	1902	7	US-10-424-599-100253	Sequence 20520, A
14	535	26.3	1644	7	US-10-425-114-20520	Sequence 849, App
15	527.5	25.9	1130	6	US-10-225-066A-849	Sequence 319, App
16	527.5	25.9	1130	9	US-10-374-780A-319	Sequence 849, App
17	527.5	25.9	1130	9	US-10-225-066A-849	Sequence 101230, A
18	526	25.9	390	8	US-10-425-115-101230	Sequence 121702, A
19	525.5	25.8	1319	8	US-10-425-115-121702	Sequence 287, App
20	518	25.5	1257	6	US-10-225-066A-287	Sequence 287, App
21	518	25.5	1257	7	US-10-374-780A-2591	Sequence 2591, App
22	518	25.5	1257	9	US-10-225-066A-287	Sequence 23359, A
23	517.5	25.4	1152	7	US-10-437-963-25359	Sequence 844, App
24	517	25.4	1876	7	US-10-374-780A-844	Sequence 20659, A
25	516	25.4	1580	7	US-10-425-114-20659	Sequence 12858, A
26	514	25.3	1032	7	US-10-425-114-12858	Sequence 44069, A
27	509	25.0	740	7	US-10-424-599-44069	Sequence 6, Appli
28	509	25.0	1209	9	US-10-737-164-6	Sequence 807, App
29	508.5	25.0	1125	6	US-10-225-066A-807	Sequence 97, Appl
30	508.5	25.0	1125	7	US-10-374-780A-97	Sequence 807, App
31	508.5	25.0	1125	9	US-10-225-066A-807	Sequence 2607, App
32	508	25.0	894	7	US-10-374-780A-2607	Sequence 879, App
33	508	25.0	894	7	US-10-412-699B-879	Sequence 145, App
34	508	25.0	894	8	US-10-495-918-145	Sequence 73, Appl
35	508	25.0	1260	6	US-10-278-536-73	Sequence 1035, App
36	508	25.0	1260	6	US-10-225-066A-1035	Sequence 177, App
37	508	25.0	1260	9	US-10-302-267-177	Sequence 1035, App
38	508	25.0	1260	9	US-10-225-066A-1035	Sequence 838, App
39	507.5	25.0	1886	7	US-10-374-780A-838	Sequence 2167, App
40	507	24.9	702	7	US-10-437-963-2167	Sequence 5337, App
41	505	24.8	1851	7	US-10-425-114-5337	Sequence 137682, A
42	505	24.8	1811	8	US-10-425-115-137682	Sequence 14719, A
43	503	24.7	1352	7	US-10-425-114-14719	Sequence 29, Appl
44	503	24.7	1324	5	US-10-286-264-29	Sequence 175, App
45	503	24.7	1324	6	US-10-278-536-175	

ALIGNMENTS

RESULT 1
US-10-780-703-1

- ; Sequence 1, Application US/10780703
- ; Publication No. US20050034194A1
- ; GENERAL INFORMATION:
- ; APPLICANT: Lee, Jong Seob
- ; APPLICANT: Kim, Yun Hee
- ; APPLICANT: Choi, Eun kyung
- ; APPLICANT: Yoo, So yeon
- ; APPLICANT: Ahn, Ji Hoon
- ; APPLICANT: Choi, Yang Do
- ; TITLE OF INVENTION: Gene Controlling Flowering Time of Plants and Method for
- ; TITLE OF INVENTION: Manipulating Flowering Time of Plant Using the Same
- ; FILE REFERENCE: 012679-105
- ; CURRENT APPLICATION NUMBER: US/10780, 703
- ; CURRENT FILING DATE: 2004-02-19
- ; PRIOR APPLICATION NUMBER: KR 10-2003-10772
- ; PRIOR FILING DATE: 2003-02-20
- ; NUMBER OF SEQ ID NOS: 9
- ; SOFTWARE: FastSeq for Windows Version 4.0
- ; SEQ ID NO 1
- ; LENGTH: 1140
- ; TYPE: DNA
- ; ORGANISM: Arabidopsis thaliana
- ; FEATURE:
- ; NAME/KEY: CDS
- ; LOCATION: (1)...(1137)
- ; OTHER INFORMATION: CDNA of LOV1 gene
- US-10-780-703-1

Alignment Scores:

Pred. No.:	7.34e-220	Length:	1140
Score:	2034.00	Matches:	379
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0
US-10-780-703-2 (1-379) x US-10-780-703-1 (1-1140)			
Qy	1	MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn	20
Db	1	ATGCAATTGTATCTCTCACAAACAGCATCATTCCTCATGAGTAACTCAAGTCAACAATAAC	60
Qy	21	GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu	40
Db	61	GAATAAGGTATAGAAAGCAATCATCATAGAGCGGCCAAGAGAGTCATGTCCAAATGAA	120
Qy	41	AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr	60
Db	121	GATGAGCTGATGATCATCATCATGACATGTCATGTCGCGGATTTAGATTCATCTCTACC	180
Qy	61	GluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal	80
Db	181	GAAGAAGAACTCATAGAGTTTACCTTCGCCGAAAGTTGAAGGCAACCGCTTTAATGTA	240
Qy	81	GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProIleAla	100
Db	241	GAACTCATCATCTTCCTCGATCTTTATCGCTATGATCTCTTGGGAATCTTCTCTCTATGGCG	300
Qy	101	AlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGly	120
Db	301	CGCATAGGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAATATAGAAATGGA	360
Qy	121	AspArgProAsnArgValThrThrSerGlyTyrTyrLysAlaThrGlyAlaAspArgMet	140
Db	361	GATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAAGCCACCGAGCTCATAGGATG	420
Qy	141	IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly	160
Db	421	ATCAGATCGGAGACTTCTCGGCTATCGGATTTAAAGAAACCCCTAGTTTCTACTCTGT	480
Qy	161	LysAlaProLysGlyThrArgThrSerTyrIleMetAsnGluTyrArgLeuProHisHis	180
Db	481	AAAGCCCTTAAAGGCATCGTACTAGTTGGATCATGACGAGTATCGTCTTCGACCAT	540
Qy	181	GluThrGluLysTyrGlnLysAlaGluLeSerLeuCysArgValTyrLysArgProGly	200
Db	541	GAACCCGAGAAGTACCAAAAGGCTGAAATATCATTTGTGCGGAGTGTACAAAAGGCCAGA	600
Qy	201	ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer	220
Db	601	GTAGAAGATCATCCATCCGTACCGACTTCTCTCTCCAAAGACATCATTAACCATTA	660
Qy	221	SerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerAsnHis	240
Db	661	TCGACATCATCCGTTTAGCCTTAAGACAAACAAACACCATTCCTCTCTCTAATCAT	720
Qy	241	SerAspAsnAsnLeuAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu	260
Db	721	TCCGACAAACACTTTAAACAAACAAACAAACATCAACATCTCGAGAAGCTCTCCACCGAA	780
Qy	261	TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThr	280
Db	781	TATTCCGCGGACGGCAGCAACAAACAGCACCAACAAACAGTAATCTCTGACGTGTACAT	840
Qy	281	AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu	300
Db	841	GCTCTAGCAATCAAAACATATATCGTCCATGCTTACGACACACAAACAAACATG	900
Qy	301	IleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValAspAspLeuGln	320
Db	901	ATAGTCTCTACGAGAAATCATCAAGACGATGATGAACACTGCCATTTGTCACATCTTCAA	960

448	Db	CGCATAGGAGAGAAAGTGGTACTTCTATGTGCCAAGAGATCGGAATAATATAGAAATGGA	507
121	Qy	AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet	140
508	Db	GATAGACCGAACCGAGTTACGACTTCAGGATATTGGAAAGCCACCGAGCTGATAGGATG	567
141	Qy	ILeArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly	160
568	Db	ATCAGATCGGAGACTTCTCGGCCTATCGGATTTAAAGAAAAACCTTAGTGTTCCTACTCTGGT	627
161	Qy	LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis	180
628	Db	AAAGCCCTTAAAGGCATCGTACTAGTGTGATCATGAACGAGTATCGTCTTCGCGACCAT	687
181	Qy	GluThrGluLysTrpGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGly	200
688	Db	GAACCCGAGAGTACCAAAAGGCTGAATATCATTTGTGCCGAGTGTTACAAAAGGCCAGGA	747
201	Qy	ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer	220
748	Db	GTAGAAGATCATCCATCGGTACCAAGTCTCTCTCCACAAGACATCAATACCATAACTCA	807
221	Qy	SerThrSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerSerAsnHis	240
808	Db	TCGACATCATCCCGCTTTCAGCCTTAAGACAAACAACACCATTCATCCTCCTCTAATCAT	867
241	Qy	SerAspAsnAsnLeuAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu	260
868	Db	TCCGACAAACAACTTAAACAACAACAACATCAACAATCTTCGAGAAGCTCTCCACCGAA	927
261	Qy	TyrSerGlyAspGlySerThrThrThrThrThrAsnSerAsnSerAspValThrIle	280
928	Db	TATTCGGCGACGGCAGCACACAACAACGACCAACAACAGTAACCTCTGACGTTTACCATT	987
281	Qy	AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu	300
988	Db	GCTCTAGCCAAATCAAAACATATATCGTCCAAATGCTTACGACACAACACAACACATG	1047
301	Qy	IleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValAspAspLeuGln	320
1048	Db	ATAGTCTCTACGAGAAATCATCAAGACGATGAATGCCATTGTTCGACATCTTCAA	1107
321	Qy	ArgLeuValAsnTyrGlnIleSerAspGly-----	330
1108	Db	AGACTAGTTTAATACCAAAATATCAGATGGAGGTAAACATCAATCACCATACTTTCAAATT	1167
330	Qy	-----	330
1168	Db	GCTCAACAGTTTTCATCATCTCAACAACAAAAATGCTAACGCAACGCAATTACAATTCGGTG	1227
331	Qy	-----AlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn	345
1228	Db	GCTGCGCGACTACAGCGCAACGCTAATGCCTCAAACTCAAGCGGGTGTAGCTATGAAC	1287
346	Qy	MetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleVal	365
1288	Db	ATGATTCCTGCAGAGACGATTCCAAACAAATGCTTTGTGGGNTATGTGGAAATCCAATAGTA	1347
366	Qy	ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys	379
1348	Db	CCAGATGGAAACAGAGATCACTATCAATATTCCTTTAAG	1389

RESULT 4

RESULT 4	
US-10-374-780A-2475	
; Sequence 2475, Application US/10374780A	
; Publication No. US20040019927A1	
GENERAL INFORMATION:	
; APPLICANT: Sherman, Bradley K	
; APPLICANT: Riechmann, Jose Luis	
; APPLICANT: Jiang, Cai-Zhong	
; APPLICANT: Heard, Jacqueline E	
; APPLICANT: Haake, Volker	
; APPLICANT: Creelman, Robert A	

61	GluGluGluLeuLeuIleGluPheTyrIleuAargArgIysValGluGlyLysAargPheasnVal	80
328	GAAGAAGAACTCATAGAGTTTTACCTTCGCCGAAAGTTTGAAGGCAACAGCTTTTAATGTA	387
81	GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProalaMetala	100
388	GAACTCATCACTTCTTCGATCTTTATCGTATGATCCTTGGGAACTTCTCTCTATGGCG	447
101	AlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgangly	120

448	Db	CGCATAGGAGAGAAAGAGCTGGTACTTCTATGTGCGCAAGAGATCGGAAATATAGAAATGGA	507
121	Qy	AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyValaAspArgMet	140
508	Db	GATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAAGCCACCGGAGCTGATAGGATG	567
141	Qy	IleArgSerGluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSerGly	160
568	Db	ATCAGATCGGAGACTTCCTCGGCCCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGGT	627
161	Qy	LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis	180
628	Db	AAAGCCCTAAAGGCACTCGTAGTTCAGTATGAACGAGATATCGTCTCCGACCAT	687
181	Qy	GluThrGluLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGly	200
688	Db	GAAACCGAGAAAGTACCAAAAGGCTGAAATATCATTTGTGCCGAGTGTACAAAGGCCAGGA	747
201	Qy	ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer	220
748	Db	GTAGAAGATCATCCATCGTACCAAGTCTCTCTCCACAAGACATCATAAACATAACTCA	807
221	Qy	SerThrSerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSerSerSerAsnHis	240
808	Db	TCGACATCATCCCCGTTTAGCCCTTAAGACAAACAACACCACTCATCTCTCTTAATCAT	867
241	Qy	SerAspAsnAsnLeuAsnAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu	260
868	Db	TCCGACAAACACTTAAACAACAACAACATCAACAATCTCGAGAGAGCTCTCCACCGAA	927
261	Qy	TyrSerGlyAspGlySerThrThrThrThrThrThrAsnSerAsnSerAspValThrIle	280
928	Db	TATTCCGGCGAGCGGCAGCACAAACAACGACCAACAACAGTAATCTCTGACGTTTACCATT	987
281	Qy	AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu	300
988	Db	GCTCTAGCCAAATCAAAACATATATCGTCCATATGCTTACGACACAAGCAACAACACATTG	1047
301	Qy	IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGln	320
1048	Db	ATAGTCTCTACGAGAATCATCAAGACGATGATGAACTGCCATTGTTGACATCTTCAA	1107
321	Qy	ArgLeuValAsnTyrGlnIleSerAspGly	330
1108	Db	AGACTAGTTTAACCTACCAAAATATCAGATGGAGGTAACATCAATCACCATACTTTCAAATT	1167
330	Qy	-----	330
1168	Db	GCTCAACAGTTTTCATCATCTACTCAACAACAANAATGCTAACGCAAAACGCATTACAAATTGGTG	1227
331	Qy	-----AlaThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn	345
1228	Db	GTCGGGGGACTACAGCGCACAAACGCTAATGCCTCAAACTCAAGCGGGCTTAGCTATGAAC	1287
346	Qy	MetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleVal	365
1288	Db	ATGATTCTCGAGGAACGATTCCAACAACATGCTTTGTGGGATATGTGGAAATCCAATAGTA	1347
366	Qy	ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys	379
1348	Db	CCAGATGGAAACAGAGATCACTACTAATAATTCCTTTAAG	1389

RESULT 5

RESULT 5	
US-10-225-066A-115	
; Sequence 115, Application US/10225066A	
; Publication No. US20050160493A9	
; GENERAL INFORMATION:	
; APPLICANT: Mendel Biotechnology, Inc.	
; APPLICANT: RATCLIFFE, Oliver	
; APPLICANT: RIECHMANN, Jose Luis	
; APPLICANT: ADAM, Luc J	
; APPLICANT: DUBELL, Arnold T	
; APPLICANT: HEARD, Jacqueline E	
<hr/>	
Qy	121 AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet 140
Dd	508 GATAGACCGAACCCGACTTCAGCATATTGGAAAGCACCACCGAGCTGTAGTAGTG 567
Qy	141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
Dd	568 ATCAGATCGAGACTTCTCGGCCTATCGGATTAAAGAAAACCCCTAGTTTCTACTCTGGT 627
Qy	161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGlyTyrArgLeuProHisHis 180

```
Db 628 AAAGCCCTAAAGCGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCTTCGCGACCAT 687
QY 181 GluThrGluLysTyrGlnLysAlaGluIleSerLeuCyshArgValTyrLysArgProGly 200
Db 688 GAAACCGAGAGTACCAAAAGCGTGAATATCATTTGTCGCGAGTGTACAAAAGCGCAGGA 747
QY 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnSer 220
Db 748 GTAGAGATCATCATCGGTACCACTCTCTCTCCACAGACATCATTAACCACTACTCA 807
QY 221 SerThrSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerAsnHis 240
Db 808 TCACATCATCCCGTTAGCGTTAAGACAAACAACACCATTCATCTCTCTCTAATCAT 867
QY 241 SerAspAsnLeuAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu 260
Db 868 TCCGACAAACACCTTAACAACAACAACATCAACATCTCGAAGCTCTCCACCGAA 927
QY 261 TyrSerGlyAspGlySerThrThrThrThrThrThrAsnSerAspValThrIle 280
Db 928 TATTCGGCGACGGCAGCACACACACACACACACACACACACACACACACAT 987
QY 281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnThrLeu 300
Db 988 GCTTAGCCAATCAAAACATATATCGTCCAAATGCTTACGACACAAGCAACACACATTG 1047
QY 301 IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGln 320
Db 1048 ATAGTCTCTACGAGAATCATCAAGACGATGATGAACTGCCAATGTTGACGATCTTCAA 1107
QY 321 ArgLeuValAsnTyrGlnIleSerAspGly- 330
Db 1108 AGACTAGTTAACTACCAATATCAGATGGAGGTAAACATCAATCAACCAATCTTCAAAT 1167
QY 330 ----- 330
Db 1168 GCTCAACAGTTTTCATCATCTCAACAACAAATGCTAAACGCAACGCAATTACAAATTTGGTG 1227
QY 331 -----AlaThrThrLeuMetProGlnThrGlnAlaIleAlaLeuAlaMetAsn 345
Db 1228 GCTCGCGCGCACTACGCGACACACGCTAAATGCTCAAACTCAACGCGGTTAGCTATGNAAC 1287
QY 346 MetIleProAlaGlyThrIleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleVal 365
Db 1288 ATGATTCCTCGAGAACGATTCACAAACATGCTTTGGGATATGGAAATCCCAATAGTA 1347
QY 366 ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
Db 1348 CCAGATGGAAACAGAGATCACTATCTATTAATTTCTTTTAAAG 1389
```

RESULT 6

```
US-10-780-703-3
; Sequence 3, Application US/10780703
; Publication No. US20050034194A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Jong Seob
; APPLICANT: Kim, Yun Hee
; APPLICANT: Choi, Eun Kyung
; APPLICANT: Yoo, So Yeon
; APPLICANT: Ahn, Ji Hoon
; APPLICANT: Choi, Yang Do
; TITLE OF INVENTION: Gene Controlling Flowering Time of Plants and Method for
; FILE OF INVENTION: Manipulating Flowering Time of Plant Using the Same
; FILE REFERENCE: 012679-105
; CURRENT APPLICATION NUMBER: US/10780, 703
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: KR 10-2003-10772
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2606
; TYPE: DNA
```

```
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)....(2606)
; OTHER INFORMATION: genomic DNA of LOV1 gene
US-10-780-703-3
```

```
Alignment Scores:
Pred. No.: 1,968-185 Length: 2606
Score: 1735.00 Matches: 378
Percent Similarity: 43.5% Conservative: 0
Best Local Similarity: 43.5% Mismatches: 1
Query Match: 85.3% Indels: 1
DB: 8 Gaps: 3
```

US-10-780-703-2 (1-379) x US-10-780-703-3 (1-2606)

```
QY 1 MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
Db 1 ATGGCAATTTGTATCTCTCCACAAAGCATCAITCCCATGAGTAACCAAGTCAACAATAAC 60
QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
Db 61 GAAAAAGGTATAGAGACAATGATCATAGAGCGGCCAAGAGAGTCATGTCCAAAATGAA 120
QY 41 AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr 60
Db 121 GATGAAGCTGATGATCATGATCATGATGTCATGTCATGTCGCGGATTTAGATTCCTACTACC 180
QY 61 GluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
Db 181 GAAGAAGAACTCATAGAGTTTTACCTTCGCCGAAAGTTGAAGGCAAAAGCTTTAATGTA 240
QY 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuPro- 97
Db 241 GAACATCATCACTTCTCGATCTTTATCGTATGATCCTTGGAACTTCTCGTAAATAT 300
QY 97 ----- 97
Db 301 ACATTCACATAAACAACACATAAATCATCTCAAACTATTGGAAATCTTAATTTCTATTCA 360
QY 97 ----- 97
Db 361 TATGTTAAGATCTTCTCTCTCTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
QY 97 ----- 97
Db 421 ATATATGTACCTACCTCCTTATGAAGTATTACTATGTGATCGATCGTTAACTTCTCAATAT 480
QY 97 ----- 97
Db 481 CTTTAAACGCTTCTCCCTCTTTAGTTTCTTTCTTAAATTAACCTAATTAACAACCTACA 540
QY 97 ----- 97
Db 541 TATATATCATAGATATACAAATATGTTATGTTCTTTTATATAGCTTATGTTATGTTAA 600
QY 98 -----AlaMetAlaIleGlyGluLysGluTyrPheTyr 110
Db 601 TCATAGATATATGATATGTCAGCTATGCGCGCATAGGAGAGAGAGTGTACTTCTAT 660
QY 111 ValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGly 130
Db 661 GTGCCAAGAGATCGGAAATATAGAAATGGAGATAGACCGAACCGAGTACGACTTCAGGA 720
QY 131 TyrTyrLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGly 150
Db 721 TATTGMAAGCCACCGAGCTGATAGATGATCAGATCGAGACTTCTCGGCTATCGGA 780
QY 151 LeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyr 170
Db 781 TTAAGAAACCCCTAGTTTCTTACTCTGTTAAAGCCCTTAAAGGCACTCGTACTAGTTGG 840
```

171 IleMetAsnGluThrArgLeuProHisGluThrGluLysValGlnLysala- 188
 841 ATCATGAACGAGTATCGTCTCCGACCATGAACCGAGAGTACCAAAAGGT-ATAAAT 899
 188 ----- 188
 900 TCTACTATACTTATATATATATCTATTTATACATACATAGATATAACCCCTAGCTAGGTG 959
 188 ----- 188
 960 GTGAGGCTTTAAATTTGAATTAATCCCTAGACAGTTTGAATTTTTCTTTTGGACTA 1019
 188 ----- 188
 1020 GTTTTATTTATTTTGGAAATTTGATTCGATAAGATCAAAAATACTTGTGATGGACTA 1079
 188 ----- 188
 1080 AATGTCAGCGCGCTTTTCGCTTAAATCCAGAAAAATTTTCATGTCATATGCGTGAATC 1139
 188 ----- 188
 1140 TTTAAATTTGTAGACATGGCCCATATGTTTATAGTAGAATACATTAATAGATAGATGATA 1199
 188 ----- 188
 1200 CACATATATAACACACAGATATCACATCGACATTCATATACCTTAATTCGACAGAG 1259
 188 ----- 188
 1260 ACATAGTTAGTTTCTTACAATTTATGACATGAATGTTCTGCTCTTCTCTCACATTAAT 1319
 188 ----- 188
 1320 TCATGCTCTTCTATTAAGTTTACCAACATTTTGTGAATAAATTTGGCATATATGAATAT 1379
 188 ----- 188
 1380 ACCAATATTTATATGCGAACATTTAAATCTATACGAATGATAACGGTTTATGGAGTA 1439
 188 ----- 188
 1440 GACCGAAAAATATATGATATACGAAAAATGACAAATGATAGATAAATACATTTTGGG 1499
 188 ----- 188
 1500 CTCTTCGACTTATATGTCGTCAACATTTGAACCAATAAATTTATAAAATTTTCTATGTA 1559
 188 ----- 188
 1560 TATATATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
 188 ----- 188
 1620 CTATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
 188 ----- 188
 1680 TATAAAAAATAAATAAT 1739
 188 ----- 188
 1740 TACTCTAAATTTCTTAAAGGCTGCTATCTTTTCTGCTAAATGCTAACTTTACATAGTTT 1799
 188 ----- 188
 1800 GTGAATCTTCTTCAAAACCATATCTTCGATAAATGATATTTTTCATAGATATTTGTTAGT 1859
 188 ----- 188
 1860 CTATATTTGATAATTTGATATATGATCAAGTCTCTAAATCAATGCTGCTCATGATATAATTA 1919
 189 -----GluIleSerLeuCysArgValTyrLysArgProGlyValGluAspHisProSer 206

1920 TAGGCTGAATATATCATTTGTGCCAGTGTACAAAAGGCCAGGAGTAGAAGATCATCCATCG 1979
 207 ValProArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeu 226
 1980 GTACCAAGCTTCTCTCTCCACAGACATCATAAACATTAATTCCTGCGGCGAGCGAGC 2039
 227 AlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSer 246
 2040 GCCTTAAGACACACACACACACATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2099
 247 AsnAsnAsnAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 266
 2100 AACAAACAAACATCAACATCTCGAGAAGCTCTCCACCGAATATTCCTGCGGCGAGCGAGC 2159
 267 Thr 286
 2160 ACAACAAACACGACCAACACGATTAATCTGACGTTACCAATTCCTGACCAATCAAAAC 2219
 287 IleTyrArgProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsn 306
 2220 ATATATCGTCCATGCTTACGACACACACACACACACATTTGATAGTCTCTCTACGAGAA 2279
 307 HisGlnAspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGln 326
 2280 CATCAAGACGATGATGAACCTGCACTGTTGACGATCTTCAAAGACTAGTTAACTACCAA 2339
 327 IleSerAspGly----- 330
 2340 ATATCAGATGGAGGTAACATCAATCAACATATCTTTCAAATTCGTCAACAGTTTTCATCAT 2399
 331 -----Ala 331
 2400 ACTCAACAAACAAATGCTTAACGCAACGCAATTAACAATTTGCTGCGGCGAGCTACAGCG 2459
 332 ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThr 351
 2460 ACAACGCTAATGCTCAAACTCAAGCGGCTTAGCTATGAACATGATTTCTGCGAGGAACG 2519
 352 IleProAsnAsnAlaLeuThrAspMetTrpAsnProIleValProAspGlyAsnArgAsp 371
 2520 ATTCCAAACAAATGCTTTGTTGGATATGTGGAATCCAATAGTACCAGATGGAACAGAGAT 2579
 372 HisTyrThrAsnIleProPheLys 379
 2580 CACTATAATAATATCTTTTAAG 2603

RESULT 7

US-10-424-599-121054
 ; Sequence 121054, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic David J
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 121054
 ; LENGTH: 785
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_80317C.1
 ; US-10-424-599-121054

Alignment Scores: 2.63e-90 Length: 785
 Pred. No.: 889.00 Matches: 171
 Score:

APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 81769
LENGTH: 743
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(743)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_81262C.1
US-10-437-963-81769

Alignment Scores:
Pred. No.: 4,23e-78 Length: 743
Score: 781.00 Matches: 139
Percent Similarity: 89.5% Conservative: 15
Best Local Similarity: 80.8% Mismatches: 18
Query Match: 38.4% Indels: 0
DB: 7 Gaps: 0

US-10-780-703-2 (1-379) x US-10-437-963-81769 (1-743)

QY 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGlu 63
DB 178 GACGGCAGCGCAGACATGTTGATGTCGGGGTTCGGTTCCACCGCAGGAGGAG 237
QY 64 LeuileGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuile 83
DB 238 CTCATCGAGTTCTACCTCCGGCGAAGGTGAGGCGAAGCGTTCAACATCGAGCTCATC 297
QY 84 ThrPheLeuAspLeuThrArgTyArgTyArgProThrGluLeuProAlaMetAlaileGly 103
DB 298 GCCTTCGTGACCTCTACCGCTATGACCGGTGGGATCTTCCCGCTCTGCGCTCAATTGGG 357
QY 104 GluLysGluThrPheThrValProArgAspArgLysTyArgAsnGlyAspArgPro 123
DB 358 GACNAGAGTGTTCTTCTATGTGCCNAGACCGCATGACGAAACCGCGATCGGCC 417
QY 124 AsnArgValThrThrSerGlyTyTrpLysAlaThrGlyAlaAspArgMetileArgSer 143
DB 418 AACCGAGTGACGCATCGGGGTACTGGAAGGCCACAGGGCGGATAGGATGGTGAAGTG 477
QY 144 GluThrSerArgProLleGlyLeuLysLysThrLeuValPheTySerGlyLysAlaPro 163
DB 478 GAGGGTGACCGCTCTATCGGCCTCAAGAACAGCCTGCTTCTACGTCCGCAAGGCGCC 537
QY 164 LysGlyThrArgThrSerTrpIleMetAsnGluTyArgLeuProHisGluThrGlu 183
DB 538 AAGGGGCTTCGACAGCTGGATCATGACAGTATGCTCTCTCTATGCGACGCTGAC 597
QY 184 LysTyGlnLysAlaGluileSerLeuLysArgValTyLysArgProGlyValGluAsp 203
DB 598 CGCTACCAAAAGGTAGAAATTCGTGTGTCGAGTCTACAAACGCCAGGAATCGAAGAC 657
QY 204 HisProSerValProArgSerLeuSerThrArgHis 215
DB 658 AACTGTAGTCGATCGAATCTCGATCTACCCGCTAC 693

RESULT 11

US-10-425-115-173486
Sequence 173486, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 173486
LENGTH: 792
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_897C.1
US-10-425-115-173486

Alignment Scores:
Pred. No.: 4,52e-73 Length: 792
Score: 737.00 Matches: 139
Percent Similarity: 80.0% Conservative: 13
Best Local Similarity: 73.2% Mismatches: 22
Query Match: 36.2% Indels: 16
DB: 8 Gaps: 2

US-10-780-703-2 (1-379) x US-10-425-115-173486 (1-792)

QY 41 AspGluAlaAspAspHisAspMetValMetProGlyPheArgPheHisProThr 60
DB 250 GAGCGCGGGACGCGCAGCACACGCTGGTGTGATGTCGGGGTTCGGTTCCACCCACG 309
QY 61 GluGluGluLeuileGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
DB 310 GAGGAGGAGTGATCGAGTTCTACCTCCGGCGCAAGTTGAGGGAAACGCTTCAACGTC 369
QY 81 GluLeuileThrPheLeuAspLeuThrArgTyArgTyArgProThrGluLeuProAlaMetAla 100
DB 370 GAGCTCATTTGCTTCTCGACCTCTACCGCTTCGACCCATGGGAGCTTCCAGCAATGGCG 429
QY 101 AlaileGlyGluLysGluThrPheThrValProArgAspArgLysTyArgAsnGly 120
DB 430 GTGATGGCGGGAGAGAGTGTTCTTCTACGTCCGAGGACCGTAACTACCGAACGGA 489
QY 121 AspArgProAsnArgValThrThrSerGlyTyTrpLysAlaThrGlyAlaAspArgMet 140
DB 490 GACCCGCCAAACCGGGTGACGGCGTACTGGAAGGCCAGCGCGCTGACCGGATG 549
QY 141 IleArgSerGluThrSerArgProLleGlyLeuLysLysThrLeuValPheTySerGly 160
DB 550 ATCCGAGGCGAGAACAAACCGCCCTCGGGCTGGAAGAGACGCTCGTGTCTACTCCGCG 609
QY 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyArgLeuProHisHis 180
DB 610 AAGGGCGCCCAAGGGCGTCCGACAGCTGGATCATGACAGTATCGCTTCCGCCACCA 659
QY 181 GluThrGluLys-----TyGlnLysAlaGluileSerLeuLysArgValTyLys 197
DB 670 CCCACGCGCGCGCCCATTTGATTCCTCCAGTCCGAGATCTCTCTGCGCGCTCTACAAG 729
QY 198 ArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsn 217
DB 730 CGCTCAGGCATCGACGAC-----GGC 750
QY 218 HisAsnSerSerThrSerArgLeuAla 227
DB 751 CACGGCGAGTCTCTCTCGACGACCCCAAGCT 780

RESULT 12

US-10-437-963-53771/C
Sequence 53771, Application US/10437963
Publication No. US2004012343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.


```

QY 113 gAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 133
Db 469 AGAAGAGAGTACCCAAAGTGGCGAGGCCAAACAGCGCTGCAACTTACGGGTATTGGAA 528
QY 133 sAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 153
Db 529 AGCCACGGGTACAGATAAGCCATATATGTAGTGAACCTCAAAAG---GTTGGGGTGAAGAA 585
QY 153 sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAs 173
Db 586 ATCTTTGGTTTCTATGGAGGAAGCCACCAAAAGGGGTCAAAACTGATTGGATCATGCA 645
QY 173 nGluTyrArgLeuProHis-----Hi 180
Db 646 TGAGTACCGGTGTACTGAAACCAAGCCTAACATAGCGCTCTGGGTGTGACTTGGGCCA 705
QY 180 sGluThrGlyLysTyrGlnLysAlaGluIleSerLeuCyAsArgValTyrLysArgPro 200
Db 706 CAAGAAAACCTCCCTAAGGCTGGATGATTGGGTGTTGTCCGGATCTACAAGAGGCCAA 765
QY 200 yValGlu---AspHisPro-----SerValProAr 209
Db 766 CACACAAAGGTACATGACAGAGGGATGATTCATGGATGATGATGGAGAAGTACTCTCC 825
QY 209 gSer-----LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSe 224
Db 826 TTCCATCAACAGTGGGCCACATGATGATGGGTGTTGTCCCTTCAAGATGTCACAGAGCTA 885
QY 224 rArgLeuAlaLeuArgGlnGlnHisHisSerSerSer-----SerAsnHi 240
Db 886 CAGTGTGTGATTTTGGAAATACACAGAAACACGTTAGAGGGGTGTTATAGGCAATGG 945
QY 240 sSerAspAsnAsnLeuAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrG 260
Db 946 TAGTGTGAATGGATCAGCAATACTAATAATATACAAAGTGTCAATCATCATCTCATCA 1005
QY 260 uTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThr 280
Db 1006 GTTTGGC-----ACCTCAAACTCCAAAGGCAGAGCT-----1036
QY 280 eAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThr 300
Db 1037 -----CTTCCTTTGTCCCTTCAACACACACTTC 1065
QY 300 uIleValSerThrArg-----AsnHi 307
Db 1066 TAATTGGCTCCAAAGAACACTCTCATCACTCTATTGGAATGTTGATGATGATAATGA 1125
QY 307 sGlnAspAsp-----AspGluThrAlaIleValAspAspLeuGlnAr 321
Db 1126 TGATGATGATACAAACACTTCAATTGGACAGATAATGGAAATGTGAGTGTGTTGAGGAC 1185
QY 321 gLeuValAsnTyrGlnIleSerAspGly-----AlaThrThrLeu-----334
Db 1186 TGNAGAGATAATAATAATACTAATGGTACTTCTGGCTCTTTGGCCACTCTGCTTAACCA 1245
QY 335 -MetProGlnThrGlnAlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAs 354
Db 1246 GCTTCACAAACACTTCAATGTCGCCAAATAGATGG-GTTACTTCCGAACCATATCAAA 1304
QY 354 nAsnAlaLeuTrpAspMetTrp 361
Db 1305 TACAAGGACCAATGGTATGG 1326

```

RESULT 14

```

US-10-425-114-20520
; Sequence 20520, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.

```

```

; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21153313/B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20520
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3242-006-H7_FLI
US-10-425-114-20520

```

```

Alignment Scores:
Pred. No.: 1,06e-49 Length: 1644
Score: 535.00 Matches: 129
Percent Similarity: 50.9% Conservative: 73
Best Local Similarity: 32.5% Mismatches: 130
Query Match: 26.3% Indels: 66
DB: 7 Gaps: 12

```

US-10-780-703-2 (1-379) x US-10-425-114-20520 (1-1644)

```

QY 14 SerAsnGlnValAsnAsnAsnGluLysGlyIleGluAspAsnAspHis-ArgGlyGly 33
Db 90 AACACATAGTAGAACAAACAAAC-----CTTGTAATAATAACCTGGAGAGCACCA 140
QY 33 nGluSerHisValGlnAsnGluAspGluAlaAspHisAspHisAspMetValMetPr 53
Db 141 AGACTCATCCAGGATTCACAGCAACACCCGAC-----CTGCCACC 182
QY 53 oGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTyrLeuArgArgLysVa 73
Db 183 CGGGTTTCGGTTCACCCACACAGAGAGCTCGTGGTTCATTACCTCAGAGAGAAAGT 242
QY 73 lGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspPr 93
Db 243 AGATTCTGTCTCTCCCTCTCAATCATCGCTGATGTTGATCTCTCAAGTTTGTATCC 302
QY 93 oTrpGluLeuProAlaMetAlaIleGlyGlyLysGluTyrPheTyrValProAr 113
Db 303 ATGGGAATTTACAGCAAGCATCTGTTGGGCGGAGAGTGTACTTTTTCAGCCCGAG 362
QY 113 gAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 133
Db 363 AGAAAGGAAGTACCCGAATGGCGAGGCCAAACAGGGCTGCACACTTCAGGGTATTGGAA 422
QY 133 sAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 153
Db 423 AGCCACGGGACAGATAAGCCATATATGTAGTGAACCTCAAAAG---GTTGGGGTGAAGAA 479
QY 153 sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAs 173
Db 480 ATCTTTGGTTTCTATGGAGGAAGCCACCAAAAGGGGTAAAAACCGATTGGATCATGCA 539
QY 173 nGluTyrArgLeuProHis-----Hi 180
Db 540 TGAGTACCGGTGTACTGAAACCAAGCCTAACATAGCGCTCTGGGTGTGACTTGGGCCA 599
QY 180 sGluThrGlyLysTyrGlnLysAlaGluIleSerLeuCyAsArgValTyrLysArgPro 200
Db 600 CAAGAAAACCTCCCTAAGGCTGGATGATTGGGTGTTGTGCCGAATCTCAAGAGAGGCCAA 659
QY 200 yValGlu---AspHisPro-----SerValProAr 209
Db 660 CACACAAAGGTACATGAGAGGAGGAGTCCATTGATGATGATGATGGAGAGTACTCTCC 719
QY 209 gSer-----LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSe 224

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2006, 09:13:07 ; Search time 1510 Seconds
(without alignments)

11502.112 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctccac.....ctaataattcctttaagtaa 2606

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2606	100.0	2606	14	ADW23849
2	2606	100.0	2606	14	ADW23849 Arabidops
3	687.2	26.4	1444	10	ADW23849 Long Vege
4	687.2	26.4	1444	10	ADW23849 Plant yie
5	687.2	26.4	1444	10	ADW23849 Plant yie
6	467.2	17.9	1140	14	ADW23847
7	467.2	17.9	1140	14	ADW23847 Arabidops
8	164	6.3	1212	8	ADA70597
9	164	6.3	1212	8	ADA70597 Rice gene
10	149	5.7	1250	8	ADA71117
11	132.6	5.1	1549	8	ADA70311
12	131.8	5.1	1563	14	ADW18167
13	120.2	4.6	824	12	ADW23847
14	115	4.4	879	12	ADW23847
15	114.2	4.4	1297	13	ADW54580
16	112.2	4.3	945	6	ABZ13738
17	112.2	4.3	945	8	ADA68463
18	112	4.3	1680	3	AAC36983
19	111.6	4.3	1635	13	ADW15275

20	111.6	4.3	1688	12	ADO62516
21	110.4	4.2	1319	10	ADD30271
22	110.4	4.2	1319	12	ADW23849
23	110.4	4.2	1319	14	ADW23849
24	110.4	4.2	1606	3	AAC42175
25	110.4	4.2	1638	3	AAC44458
26	110.4	4.2	1641	3	AAC45486
27	110.2	4.2	1359	13	ADW45355
28	110.2	4.2	1580	13	ADW45355
29	109.6	4.2	1384	12	ADW23849
30	109.6	4.2	1398	13	ADW45355
31	109.6	4.2	1557	12	ADW23849
32	108.8	4.2	1445	13	ADW23849
33	108.6	4.2	894	9	ADW23849
34	108.6	4.2	894	12	ADW23849
35	108.6	4.2	894	12	ADW23849
36	108.6	4.2	894	12	ADW23849
37	108.6	4.2	894	14	ADW23849
38	108.6	4.2	903	3	AAC43036
39	108.6	4.2	923	12	ADW23849
40	108.6	4.2	1125	10	ADW23849
41	108.6	4.2	1125	12	ADW23849
42	108.6	4.2	1125	14	ADW23849
43	108.6	4.2	1234	3	AAC50693
44	108.6	4.2	1237	3	AAC33463
45	108.6	4.2	1260	10	ADW55715

ALIGNMENTS

RESULT 1	
ADW23849	
ID	ADW23849 standard; DNA; 2606 BP.
XX	AC
XX	ADW23849;
DT	10-MAR-2005 (first entry)
XX	XX
DE	Arabidopsis thaliana LOV1 genomic sequence, SEQ ID 3.
XX	XX
KW	Flowering; plant; gene; ds.
XX	XX
OS	Arabidopsis thaliana.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..2606
FT	/tag= b
FT	/product= "LOV1 protein"
FT	1..292
FT	/tag= a
FT	/number= 1
FT	293..622
FT	/tag= c
FT	/number= 1
FT	623..892
FT	/tag= d
FT	/number= 2
FT	893..1923
FT	/tag= e
FT	/number= 2
FT	/cons_splice= (5'site:NO,3'site:NO)
FT	1924..2351
FT	/tag= f
FT	/number= 3
FT	2352..2456
FT	/tag= g
FT	/number= 3
FT	/cons_splice= (5'site:NO,3'site:NO)
FT	2457..2606
FT	/tag= h
FT	/number= 4
XX	XX

XX 29-JAN-2004 (first entry)
 XX Plant yield related polynucleotide clone G962.
 DE ds; gene; transcription factor; transgenic plant; salt stress resistance;
 XX osmotic stress resistance; freezing tolerance; drought tolerance;
 KW low humidity tolerance; radiation resistance.
 XX Arabidopsis thaliana.
 OS
 XX
 FH Key Location/Qualifiers
 FT 148..1392
 FT CDS /*tag= a
 FT /product= "transcription factor"
 XX
 PN W02003013228-A2.
 PD 20-FEB-2003.
 XX
 PF 09-AUG-2002; 2002WO-US025808.
 XX
 PR 09-AUG-2001; 2001US-0310847P.
 PR 19-NOV-2001; 2001US-0336049P.
 PR 11-DEC-2001; 2001US-0338692P.
 PR 14-JUN-2002; 2002US-00171468.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 PI Heard JE, Riechmann JL, Creelman RA, Keddle J, Pilgrim ML;
 PI Dubell AN, Jiang C, Ratcliffe O, Pineda O, Yu G, Broun PE;
 XX
 DR WPI; 2003-248222/24.
 DR P-PSDB; ADE31521.
 XX
 PT New plant transcription factor polynucleotides and polypeptides, useful
 PT in producing transgenic plants with commercially valuable properties,
 PT i.e. modified desirable traits, e.g. salt stress resistance or tolerance
 PT to freezing.
 XX
 PS Disclosure; SEQ ID NO 87; 311pp; English.
 XX
 CC The invention relates to a number of isolated cDNA sequences and their
 CC encoded proteins which are especially transcription factor related cDNA's
 CC and proteins. The isolated or recombinant plant transcription factor
 CC polynucleotides and polypeptides are useful in producing transgenic
 CC plants with commercially valuable properties, i.e. modified or altered
 CC desirable traits as compared to a reference plant, e.g. salt stress
 CC resistance, osmotic stress resistance, tolerance to freezing, drought,
 CC low humidity tolerance, or radiation resistance. Sequence information
 CC related to the polynucleotides and polypeptides can also be used in
 CC bioinformatic search methods. The transgenic plant is useful for growing
 CC a progeny plant from a parent plant. This sequence represents one of the
 CC cDNAs of the invention
 XX
 SQ Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;
 XX
 Query Match 26.4%; Score 687.2; DB 10; Length 1444;
 Best Local Similarity 96.8%; Pred. No. 2e-127;
 Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 XX 1883 GTATCAAGTCTTAATCAATGTGCTCATGTATATATATAGGCTGAATATCATCTGTGCGG 1942
 DB 669 GTATGCTTTCCGACCATGAACCGAGAGATGCAAAAGGCTGAATATCATCTGTGCGG 728
 XX 1943 AGTGTCACAAAGCCAGGAGTAGAAGATCATCCATCGGTACCAAGTCTCTCTCCACAAG 2002
 DB 729 AGTGTCACAAAGCCAGGAGTAGAAGATCATCCATCGGTACCAAGTCTCTCTCCACAAG 788
 XX 2003 ACATCATTAACCAATCAATCGATCATCCCGTTTGTAGCCCTTAAGACAAACAACACCA 2062
 DB 789 ACATCATTAACCAATCAATCGATCATCCCGTTTGTAGCCCTTAAGACAAACAACACCA 848

QY 2063 TTCTATCTCTCTAATCAATTCGCAACAACCTTAAACAACAACAACAACAATCAATCAATCT 2122
 DB 849 TTCTATCTCTCTAATCAATTCGCAACAACCTTAAACAACAACAACAACAATCAATCT 908
 QY 2123 CGAGAAGTCTCTCCACCGAATATTCGGGCGACGGGAGGCAACAACAACAACAACAACAAG 2182
 DB 909 CGAGAAGTCTCTCCACCGAATATTCGGGCGACGGGAGGCAACAACAACAACAACAACAAG 968
 QY 2183 TAACTCTGACGTTTACCATTTGCTCTAGCAATCAAAACAATATATCGTCCAAATGCTTTACGA 2242
 DB 969 TAACTCTGACGTTTACCATTTGCTCTAGCAATCAAAACAATATATCGTCCAAATGCTTTACGA 1028
 QY 2243 CACAAGCAACAACAACAATTTAGTCTCTACGAGAAATCATCAAGACGATGATCAAACTGC 2302
 DB 1029 CACAAGCAACAACAACAATTTAGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 1088
 QY 2303 CATTTGTCAGATCTTCAAAAGACTAGTTAACTACCAATATCAGATGAGGTTAACTCAA 2362
 DB 1089 CATTTGTCAGATCTTCAAAAGACTAGTTAACTACCAATATCAGATGAGGTTAACTCAA 1148
 QY 2363 TCACCAATCTTTCAAAATTTGCTCAACAGTTTCATCATCTCAACAACAACAATGCTAAACGC 2422
 DB 1149 TCACCAATCTTTCAAAATTTGCTCAACAGTTTCATCATCTCAACAACAACAATGCTAAACGC 1208
 QY 2423 AAACGCATTACAAATTTGGTGGCTGCGGCGACTACAGCGCAACGCTAAATGCTTCAAACTCA 2482
 DB 1209 AAACGCATTACAAATTTGGTGGCTGCGGCGACTACAGCGCAACGCTAAATGCTTCAAACTCA 1268
 QY 2483 AGCGCGTTAGCTATGAACATGATTTCTGCGAGAACGATTCGCAACAACAATGCTTTGTGGGA 2542
 DB 1269 AGCGCGTTAGCTATGAACATGATTTCTGCGAGAACGATTCGCAACAACAATGCTTTGTGGGA 1328
 QY 2543 TATGTGGATCCAATAGTACAGATGGAACAAGAGATCACTATCTAAATATCTTTTAA 2602
 DB 1329 TATGTGGATCCAATAGTACAGATGGAACAAGAGATCACTATCTAAATATCTTTTAA 1388
 QY 2603 GTAA 2606
 DB 1389 GTAA 1392
 XX
 RESULT 5
 ADI44012
 ID ADI44012 standard; DNA; 1444 BP.
 XX
 AC ADI44012;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Plant transcription factor related polynucleotide #1595.
 XX
 KW transgenic; plant; enhanced tolerance to abiotic stress;
 KW glyphosphate tolerance; hormone sensitivity; disease resistance;
 KW sugar sensing; flowering; flower structure; stem bifurcation;
 KW branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; ds.
 XX
 OS Unidentified.
 XX
 PN US2004019927-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 25-FEB-2003; 2003US-00374780.
 XX
 PR 18-APR-2001; 2001US-00837944.
 XX (SHER/) SHERMAN B K.
 PA (RIEC/) RIECHMANN J L.

PA (JIAN/) JIANG C.
 PA (HEAR/) HEARD J E.
 PA (HAK/) HAKE V.
 PA (CREE/) CREELMAN R A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L J.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROWN P E.
 PA (PILG/) PILGRIM M L.
 PA (DUBE/) DUBELL A N.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Brown PE;
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
 XX
 DR WPI; 2004-132245/13.
 DR P-PSDB; ADI44013.
 XX
 PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 XX
 PS Disclosure; SEQ ID NO 2475; 435pp; English.
 XX
 CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produced a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure, change in stem bifurcations, altered branching
 CC pattern, reduced apical dominance, reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins, or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This sequence represents a plant
 CC transcription factor related polynucleotide.
 XX
 SQ Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;
 Query Match 26.4%; Score 687.2; DB 12; Length 1444;
 Best Local Similarity 96.8%; Pred. No. 26-127;
 Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 Qy 1883 GTATCAAGTCTCTAATCAATGTCTCATGTATAATTATAGGCTGAATATCATTTGTGCGG 1942
 Db 669 GTATCGTCTTCGGACCATGAACCGAGAGTAGTACCAAAAGGCTGAATATCATTTGTGCGG 728
 Qy 1943 AGTGTAACAAAGGCGCAGAGTAGAAGATCATTCATCGGTACACGTTCTCTCTCCACAAG 2002
 Db 729 AGTGTAACAAAGGCGCAGAGTAGAAGATCATTCATCGGTACACGTTCTCTCTCCACAAG 788
 Qy 2003 ACATCATACCAATCAATCATCGACATCATCCGTTTAGCCTTAGACACACACACCA 2062
 Db 789 ACATCATACCAATCAATCATCGACATCATCCGTTTAGCCTTAGACACACACACCA 848
 Qy 2063 TTTCATCTCTCTAATCATTCGACAAACACCTTAAACAAACACCAACATCAATCT 2122
 Db 849 TTTCATCTCTCTAATCATTCGACAAACACCTTAAACAAACACCAACATCAATCT 908
 Qy 2123 CGAGAAGTCTCCACCGCAATATTCGGCGAGCGGAGCAACAAACACGACCAACAACAG 2182
 Db 909 CGAGAAGTCTCCACCGCAATATTCGGCGAGCGGAGCAACAAACACGACCAACAACAG 968

Qy 2183 TAACTCTGACGTACCATTTGCTCTAGCCAAATCAAAAACATATATCGTCCAAATGCCCTTACGA 2242
 Db 969 TAACTCTGACGTACCATTTGCTCTAGCCAAATCAAAAACATATATCGTCCAAATGCCCTTACGA 1028
 Qy 2243 CACAAGCAACAAACATTTGATGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 2302
 Db 1029 CACAAGCAACAAACATTTGATGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 1088
 Qy 2303 CATTTGTCAGCATCTTCAAGACGTAGTTAACTACCAATATACAGATGGAGGTAAACATCAA 2362
 Db 1089 CATTTGTCAGCATCTTCAAGACGTAGTTAACTACCAATATACAGATGGAGGTAAACATCAA 1148
 Qy 2363 TCACCAATATCTTTCAAAATTTGCTCAACAGCTTTTCATCATATCTCAACAAACAAATGCTAAAGC 2422
 Db 1149 TCACCAATATCTTTCAAAATTTGCTCAACAGCTTTTCATCATATCTCAACAAACAAATGCTAAAGC 1208
 Qy 2423 AAACGATTAACAATTTGGTGGCTGGGAGCTACAGCGACAAACGGCTTAATGCTCAAAACTCA 2482
 Db 1209 AAACGATTAACAATTTGGTGGCTGGGAGCTACAGCGACAAACGGCTTAATGCTCAAAACTCA 1268
 Qy 2483 AGCGGCGTTAGCTATGAACATGATTCCTGCGAGGAACGATTCCTCAACAAATGCTTTGTGGGA 2542
 Db 1269 AGCGGCGTTAGCTATGAACATGATTCCTGCGAGGAACGATTCCTCAACAAATGCTTTGTGGGA 1328
 Qy 2543 TATGTGGAATCCAATAGTACCAAGTGGAAACAGAGATCAGTATATCTTAATATTCCTTTTAA 2602
 Db 1329 TATGTGGAATCCAATAGTACCAAGTGGAAACAGAGATCAGTATATCTTAATATTCCTTTTAA 1388
 Qy 2603 GTAA 2606
 Db 1389 GTAA 1392
 RESULT 6
 ADW23847
 ID ADW23847 standard; cDNA; 1140 BP.
 XX
 AC ADW23847;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE Arabidopsis thaliana LOV1 coding sequence, SEQ ID 1.
 XX
 KW Flowering; plant; gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1140
 FT /*tag= a
 FT /product= "LOV1 protein"
 XX
 FN KR2004075252-A.
 XX
 PD 27-AUG-2004.
 XX
 PF 20-FEB-2003; 2003KR-00010772.
 XX
 PR 20-FEB-2003; 2003KR-00010772.
 XX
 PA (UYSE-) UNIV SEOUL NAT.
 XX
 PI Ahn JH, Choi EG, Choi YD, Kim YH, Lee JS, Yoo SY;
 XX
 DR WPI; 2005-054880/06.
 DR P-PSDB; ADW23848.
 XX
 PT New LOV1 polynucleotide encoding a polypeptide controlling the flowering
 PT time of plants, i.e., either delaying or inducing early flowering of the
 PT plants.
 XX
 PS Claim 4; SEQ ID NO 1; 25pp; Korean.
 XX

CC	The present invention relates to the novel gene LOV1 (Long Vegetative phase 1) from <i>Arabidopsis thaliana</i> and its use in controlling flowering time in plants. LOV1 inhibits the flowering stimulating gene <i>AGL20</i> . LOV1 can be used for manipulating flowering time by over-expressing LOV1 to delay flowering, or inhibiting expression of LOV1 to stimulate flowering. CC LOV1 is also useful for identifying homologous genes in other plants. The present sequence is the LOV1 coding sequence.	XX
CC	Sequence 1140 BP; 393 A; 272 C; 216 G; 259 T; 0 U; 0 Other;	
CC	Query Match 17.9%; Score 467.2; DB 14; Length 1140;	
CC	Best Local Similarity 82.3%; Pred. No. 1.7e-83;	
CC	Matches 596; Conservative 0; Mismatches 23; Indels 105; Gaps 1;	
QY	1883 GTATCAAGTCTCTAATCAATGTGCTCATGTATAATTAGGCTGAATATATCATTTGTCGCG 1942	
DB	522 GTATCGTCTCCGACCATGAACCGAGAGTACCAAAAGGCTGAATATATCATTTGTCGCG 581	
QY	1943 AGTGTCACAAAGGCGAGGTAGAAGATCATCCATCGGTACCAGTTCTCTCTCCACAAG 2002	
DB	582 AGTGTCACAAAGGCGAGGTAGAAGATCATCCATCGGTACCAGTTCTCTCTCCACAAG 641	
QY	2003 ACATCATAAACCATAACTCATCGCATCATCCCGTTTAGCCCTTAAGACAAACAAACACCA 2062	
DB	642 ACATCATAAACCATAACTCATCGCATCATCCCGTTTAGCCCTTAAGACAAACAAACACCA 701	
QY	2063 -TTTCATCTCTCTAATCATTTCCGACAACCTTTAAACAACAACATCAACAATCT 2122	
DB	702 TTTCATCTCTCTAATCATTTCCGACAACCTTTAAACAACAACATCAACAATCT 761	
QY	2123 CGAAGAGCTCTCCACCGAATATTCCGGGAGCGGAGACACACACAGGACCCACAAACAG 2182	
DB	762 CGAAGAGCTCTCCACCGAATATTCCGGGAGCGGAGACACACACAGGACCCACAAACAG 821	
QY	2183 TAACTCTGACGTTACCATTTGCTCTAGCCAAATCAAAAACATATATCGTCCAATGCCTTACGA 2242	
DB	822 TAACTCTGACGTTACCATTTGCTCTAGCCAAATCAAAAACATATATCGTCCAATGCCTTACGA 881	
QY	2243 CACAAGCAACAAACATTTGATGTCTCTACGAGAAATCATCAAGACGATGATGAACATGC 2302	
DB	882 CACAAGCAACAAACATTTGATGTCTCTACGAGAAATCATCAAGACGATGATGAACATGC 941	
QY	2303 CATTTGTTGACGATCTTCAAGACGTAGTTTAACATACCAATATCAGATCGAGGTAACATCAA 2362	
DB	942 CATTTGTTGACGATCTTCAAGACGTAGTTTAACATACCAATATCAGATCGAGGTAACATCAA 989	
QY	2363 TCACCAATACTTTTCAAATTTGCTTCAACAGTTTTTCATCATACTCAACAACAAATGCTAAACG 2422	
DB	990 ----- 989	
QY	2423 AAACGCAATTACAATTTGGTGGCTCGGGCGACTACAGGCAACACGTAATGCTCTCAACTCA 2482	
DB	990 -----AGCGCAACACGTAATGCTCTCAACTCA 1016	
QY	2483 AGCGGCGTTAGCTATGACATGATTCTCTGACGAGAACGATTCCAACAAATGCTTTGTGGGA 2542	
DB	1017 AGCGGCGTTAGCTATGACATGATTCTCTGACGAGAACGATTCCAACAAATGCTTTGTGGGA 1076	
QY	2543 TATGTGGAATCCAATAGTACCAGATGGAACACAGAGATCACTTATCTAATATTCCTTTTAA 2602	
DB	1077 TATGTGGAATCCAATAGTACCAGATGGAACACAGAGATCACTTATCTAATATTCCTTTTAA 1136	
QY	2603 GTAA 2606	
DB	1137 GTAA 1140	

[illegible]

QY 2243 CACAAGCAACACACATTGATAGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 2302
 DB 882 CACAAGCAACACACATTGATAGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 941
 QY 2303 CATTGTTGACGATCTTCAAGACTAGTTAACTACCAAAATATCAGATGGAGGTAAACATA 2362
 DB 942 CATTGTTGACGATCTTCAAGACTAGTTAACTACCAAAATATCAGATGG 989
 QY 2363 TCACCAATACTTTTCAAAATGCTCAACAGTTTCATCATCTACTCAACAAATGCTAACGC 2422
 DB 990 ----- 989
 QY 2423 AAACGCAATTACAATTGGTGGCTGCGCGGACTACAGCAACCGCTAAATGCTCAAACTCA 2482
 DB 990 -----ACGCAACCGCTAAATGCTCAAACTCA 1016
 QY 2483 AGCGCGTTAGCTATGAACATGATTCCTGCGAGAACGATTCCAACAATGCTTTGTGGGA 2542
 DB 1017 AGCGCGTTAGCTATGAACATGATTCCTGCGAGAACGATTCCAACAATGCTTTGTGGGA 1076
 QY 2543 TATGTGGAATCCAATAGTACAGATGGAAACAGAGATCACTATCTAATATTTCTTTTAA 2602
 DB 1077 TATGTGGAATCCAATAGTACAGATGGAAACAGAGATCACTATCTAATATTTCTTTTAA 1136
 QY 2603 GTAA 2606
 DB 1137 GTAA 1140

RESULT 8

ADA70597
 ID ADA70597 standard; DNA; 1212 BP.
 XX
 AC ADA70597;
 DT 20-NOV-2003 (first entry)
 DE Rice gene, SEQ ID 3920.
 KW Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX
 OS Oryza sativa.
 XX
 PN WO200300898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX
 DR WPI; 2003-175290/17.
 XX

Identifying at least one gene involved in plant resistance or response to
 pathogenic infection for conferring resistance or tolerance to a plant to
 bacterial, fungal or viral infection by determining or detecting plant
 gene expression.

Claim 6; SEQ ID NO 3920; 899pp; English.

The present invention relates to a method (M1) for identifying genes
 involved in plant resistance or response to pathogenic infection. M1
 comprises identifying a gene whose expression is significantly altered in
 the incompatible interaction of plant gene expression relative to
 expression of the gene in an uninfected plant, in a mutant plant that
 does not express a gene associated with response to pathogenic infection,
 or in a corresponding incompatible or compatible interaction. (M1) is

CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX
 SQ Sequence 1212 BP; 245 A; 421 C; 361 G; 185 T; 0 U; 0 Other;
 Query Match 6.3%; Score 164; DB 8; Length 1212;
 Best Local Similarity 70.8%; Pred. No. 6.2e-23;
 Matches 218; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 605 AGATATATGATATGAGCTATGCGGCGATAGAGAGAAAGAGTGTACTTCTATGTGC 664
 DB 161 AGCTCGATGGGGGTGCGAGCAATGCGGCGATAGGGGAGAGAGTGTCTTCTACGTGC 220
 QY 665 CAAGAGATCGGAAATATAGAAATGAGATAGACCGAACCGAGTACGACTTTCAGGATAT 724
 DB 221 CTCGGGACAGAGATACAGGAAACCGGACCGGCGAGCGGTGACGGGTGCGGGTACT 280
 QY 725 GGAAAGCCACCGGAGCTGATAGATGATCAGATCGGAGACTTCTCGGCTATCGGATTA 784
 DB 281 GGAAGCGACGGGGCGGACCGGATGATCCGAGCGGAGAACACCGCCCATCGGGCTCA 340
 QY 785 AGAAACCCCTAGTTTCTTCTACTCTGTTAAAGCCCTTAAGGCACCTGCTACTAGTTGGATCA 844
 DB 341 AGAAGACGCTTGTCTTCTACTCTCGGCAAGGCCCCCAAGGGCGTCCGACAGCTGGATCA 400
 QY 845 TGAAGAGTATCGTCTTCCGACCATGAAACCGAGAGATACCAAAAGGTATATAATTCTAC 904
 DB 401 TGAAGAGTATCGTCTTCCGACCATGAAACCGAGAGATACCAAAAGGTATATAATTCTAC 460
 QY 905 TATAACTC 912
 DB 461 TATGCCGC 468

RESULT 9

ACL26812
 ID ACL26812 standard; cDNA; 1212 BP.
 XX
 AC ACL26812;
 DT 02-JUN-2005 (first entry)
 DE Rice abiotic stress responsive polynucleotide SEQ ID NO:768.
 DE ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 KW agriculture.
 KW
 OS Oryza sativa.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 24-AUG-2001; 2001US-0314682P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 PI Moughamer T, Provart N, Ricke D, Zhu T;
 XX
 DR WPI; 2003-248011/24.
 XX

New stress-responsive nucleic acid, useful for altering the
 responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 stress, salt stress or osmotic stress.
 Claim 1; SEQ ID NO 768; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 XX
 SQ Sequence 1212 BP; 245 A; 421 C; 361 G; 185 T; 0 U; 0 Other;
 Query Match 6.3%; Score 164; DB 11; Length 1212;
 Best Local Similarity 70.8%; Pred. No. 6.2e-23;
 Matches 218; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
 QY 605 AGATATATGTATATGACGATATGCGCGCGATAGGAGAGAAAGTGGTACTTCTATGTGC 664
 DB 161 AGCTCGATGGGGTGCAGCAATGCGCGCGATAGGAGAGAAAGTGGTACTTCTATGTGC 220
 QY 665 CAAGAGATCGGAATATAGAAATGGAATAGACCGAACCGAGTAACGATTCAGGATATT 724
 DB 221 CTCGGGACAGGAATGACGAAACCGGCGCGGCGGAGTGGCGGTCTCGGGTACT 280
 QY 725 GGAAGACCCACCGAGTGTATGATGATCAGATCGGAGACTTCTCGGCTATCGGATTA 784
 DB 281 GGAAGCGGACGGGGCGGCGGCGGATGATCGAGCGGAGAACACCGCCCGATCGGCTCA 340
 QY 785 AGAAACCCCTAGTGTCTTCTACTCTGTTAAAGCCCTTAAAGGCACTCGTACTAGTTGGATCA 844
 DB 341 AGAAGACGCTTGTTCTTCTACTCTCGGCAAGGCGCCCAAGGCGTCCGACGAGCTGGATCA 400
 QY 845 TGACGAGTATCGTCTTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 904
 DB 401 TGAACGAGTATCGCTTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 460
 QY 905 TATACTC 912
 DB 461 TATGCCGC 468

RESULT 10
 ADA71117
 ID ADA71117 standard; DNA; 1250 BP.
 AC ADA71117;
 DT 20-NOV-2003 (first entry)
 XX Rice gene, SEQ ID 4440.
 DE Rice gene, SEQ ID 4440.
 XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX Oryza sativa.
 OS
 XX WO2003000898-A1.
 FN
 XX 03-JAN-2003.
 PD
 XX 22-JUN-2001; 2001WO-IB001105.
 PF
 XX 22-JUN-2001; 2001WO-IB001105.
 PR
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 DR
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 XX Claim 6; SEQ ID NO 4440; 899pp; English.
 PS
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 1250 BP; 291 A; 388 C; 339 G; 230 T; 0 U; 2 Other;
 Query Match 5.7%; Score 149; DB 8; Length 1250;
 Best Local Similarity 71.1%; Pred. No. 6.2e-20;
 Matches 197; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 QY 618 TGCAGTATGCGCGGATAGGAGAGAAAGTGGTACTTCTATGTGCCAAGAGATCGGAA 677
 DB 141 TCCCGCTCTGGCCTCAATTTGGGGACAGGAGTGGTCTTCTATGTGCCAAGGACCGCAA 200
 QY 678 ATATAGATGAGATAGACCGAACCGAGTACGACTTCAGGATATGGAAGCCACCGG 737
 DB 201 GTACCGAAACCGCGATCGGCCCAACCGAGTACGCGGTACTGGAGGCCACAGG 260
 QY 738 AGCTCATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGT 797
 DB 261 GCGGATAGGATGGTGAAGTGGAGGTGACCGTCTATCGGCTCAAGAGACGCTCGT 320
 QY 798 TTTCTACTCTGTAAAGCCCTTAAAGGCATCTGCTACTAGTTGGATCATGAAGAGTATCG 857
 DB 321 CTTCTACGTCGCAAGGCGGCCCAAGGGCTTCGCAAGCAGCTGGATCATGAAGAGTATCG 380
 QY 858 TCTTCGCGACCATGAACCGGAGTACCAAGGTA 894
 DB 381 TCTCCCTCATGGCGACGCTGACCGCTACCAAGATA 417

RESULT 11
 ADA70311
 ID ADA70311 standard; DNA; 549 BP.
 AC ADA70311;
 DT 20-NOV-2003 (first entry)
 XX Rice gene, SEQ ID 3634.
 DE Rice gene, SEQ ID 3634.
 XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX Oryza sativa.
 OS
 XX WO2003000898-A1.
 FN
 XX 03-JAN-2003.
 PD
 XX 22-JUN-2001; 2001WO-IB001105.
 PF
 XX 22-JUN-2001; 2001WO-IB001105.
 PR
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri P, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 6; SEQ ID NO 3634; 899pp; English.
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX Sequence 549 BP; 103 A; 174 C; 194 G; 78 T; 0 U; 0 Other;
 SQ
 Query Match 5.1%; Score 132.6; DB 8; Length 549;
 Best Local Similarity 69.5%; Pred. No. 9.6e-17; Mismatches 79; Indels 0; Gaps 0;
 Matches 180; Conservative 0;
 QY 620 CAGCTATGCGCGGATAGGAGAGAAAGTGTGTTCTTATGTCCAGAGATCGGAAT 679
 DB 260 CCGCAATGCGGATAGGAGAGAGAGTGTGTTCTTATGTCCGCGGACCGCAAGT 319
 QY 680 ATAGAAATGGAGATAGACCGAACCGAGTAAACGATTCAGATATTTGAAACCGCACCGAG 739
 DB 320 ACCGGAACCGGACCGCGGACCGGAGTACCGGCTCGGGTACTGGAAGCGACGCGGG 379
 QY 740 CTGATAGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCTAGTTT 799
 DB 380 CCGACCGGATGATCGGCGGAGAGAACACGCCGCCCATCGGCTCAAGAAGACGCTCGTCT 439
 QY 800 TCTACTCTGGTAAAGCCCTTAAAGCACTCCGTACTAGTTGGATCATGAACGAGTATCGTC 859
 DB 440 TCTACTCGGCAAGGCCGCCAAGGGCGTCCGACGAGCTGATCATGACGAGTACCGCT 499
 QY 860 TTCGCGACCAATGAACCGA 878
 DB 500 TCCCGCGCGCGCGCAGA 518
 RESULT 12
 ID ADW18167 standard; cDNA; 1563 BP.
 XX AC ADW18167;
 XX 24-MAR-2005 (first entry)
 XX Pinus radiata transcription factor cDNA NAC family Seq 1946.
 KW gene; ss; plant; transcription; gene regulation; gene expression;
 KW transgenic plant; drought resistance; disease resistance; salt tolerance;
 KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
 KW flower color.
 OS Pinus radiata.
 XX WO2005001050-A2.
 XX 06-JAN-2005.
 XX 07-JUN-2004; 2004WO-US017965.

XX 06-JUN-2003; 2003US-0476189P.
 XX (ARBO-) ARBORGEN LLC.
 XX Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
 PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Maguin A;
 PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
 XX WPI; 2005-075542/08.
 DR P-PSDB; ADW18488.
 XX New polynucleotides isolated from plants encoding transcription factors,
 PT and polypeptides encoded by such polynucleotides, useful for regulating
 PT gene transcription and gene expression.
 XX Claim 3; SEQ ID NO 1946; 1265pp; English.
 XX This invention relates to novel isolated plant nucleic acid molecules, or
 CC variants thereof, that encode transcription factors. Specifically, it
 CC refers to transcription factor proteins that are capable of binding to
 CC DNA in order to regulate gene transcription and gene expression in a
 CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
 CC invention describes DNA constructs containing DNA encoding a
 CC transcription factor that regulates the promoter, which is operably
 CC linked to the desired nucleic acid to be expressed. It further provides
 CC transgenic plants expressing a transcription factor that confers a trait
 CC to the plant such as increased drought, salt or disease tolerance, height
 CC change, enhanced cold/frost tolerance, enhanced color, health and
 CC nutritional characteristics, as well as improved taste, starch
 CC composition, flower longevity and germination, amongst others.
 CC Accordingly, such plants that are successfully transfected with a DNA
 CC construct can be characterized by a difference in flower color, petal or
 CC leaf shape and size, aroma or plant height. This polynucleotide is a
 CC plant transcription factor cDNA sequence of the invention.
 XX Sequence 1563 BP; 501 A; 280 C; 369 G; 413 T; 0 U; 0 Other;
 SQ
 Query Match 5.1%; Score 131.8; DB 14; Length 1563;
 Best Local Similarity 65.4%; Pred. No. 1.8e-16;
 Matches 193; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
 QY 618 TGCAGCTATGCGCGGATAGGAGAGAAAGTGTGTTCTTATGTCCAGAGATCGGAA 677
 DB 333 TCCAGCTTGGCTTCAATGAGAGAAAGTGTGTTCTTATGTCCAGATAGAGAA 392
 QY 678 ATATAGAAATGGAGATAGACCGAACCGAGTAAACGATTCAGGATATTTGAAAGCCACCGG 737
 DB 393 GTATAGGAATGGAGATCGACCTAACAGAGTTACAAAGTCTGGGTATTTGAAAGCTACAGG 452
 QY 738 AGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCTAGT 797
 DB 453 TGCTGACCGGATGTTTACAGTGAAGTCTCTGGTTGATCGGCTTGAAGAAACGCTTGT 512
 QY 798 TTTCTACTCTGGTAAAGCCCTTAAAGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857
 DB 513 GTTCTATACAGGAGGCGGCCCAAGGAAAGGACAAAGCTGGATCATGACCGAATATAG 572
 QY 858 TCTTCCGACCATGAAACCGAGAGTACCAAAAGGTATATAAATCTTACTATTAATCTC 912
 DB 573 GTTGCCCAACTCGAAACTCGCAGAAATTCAGAGAACGAACTCTCCCTTTGCCCG 627
 RESULT 13
 ID ADO62519 standard; DNA; 824 BP.
 XX AC ADO62519;
 XX 15-JUL-2004 (first entry)
 XX Transcription factor G1412 orthologous sequence, SEQ ID 986.
 XX

KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development; ds.
 XX Glycine max.
 XX WO2004031349-A2.
 XX 15-APR-2004.
 XX 18-SEP-2003; 2003WO-US030292.
 XX 18-SEP-2002; 2002US-0411837P.
 XX 17-DEC-2002; 2002US-043416P.
 XX 24-APR-2003; 2003US-0465809P.
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
 PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
 XX WPI; 2004-330163/30.
 XX New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 PT properties compared to a reference plant.
 XX Claim 1; SEQ ID NO 986; 510pp; English.

CC The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stress,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC altered seed coloration, altered seed size, altered seed shape, large
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
 CC content, altered seed protein content, altered seedprenyl content,
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and
 CC decreased anthocyanin levels. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 824 BP; 254 A; 143 C; 154 G; 272 T; 0 U; 1 Other;

Query Match 4.6%; Score 120.2; DB 12; Length 824;

Best Local Similarity 53.1%; Pred. No. 3.2e-14;

Matches 372; Conservative 0; Mismatches 289; Indels 39; Gaps 4;

QY 157 CCGGATTTCATCTCCAGGAGAGAACTCATAGAGTTTACCTTCGCGGAAA 216
 Db 164 CTTGATTTAGATTTATCCCATGATGAGGAGCTTTTGGTTAGTACCTTTGCGCGAAG 223
 QY 217 GTTGAAGGCAACGCTTTTAATGTAGAACTCATCACTTTTCTCTCGATCTTTATCGTATGAT 276
 Db 224 GTTCTGCCCATCATTTCTCTCTCCCAATCATCTGAGTTGATTTGTACAAAGTTTGTAT 283
 QY 277 CTTTGGGAACTTCTCTGGTAAAT 336
 Db 284 CCATGGGTTCTTCCAGGT-AAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 342
 QY 337 TTTGGAATCTTAAATTTCTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 396
 Db 343 ATCTTTAACCAAGATTTTGTCAATATCAATATCAATATCAATATCAATATCAATATCAATAT 402
 QY 397 TCTCTATTTCTTTTAACTTAT 456
 Db 403 GACCGTTAGATCAATTTGTCTTCAATTTCTTCCCATAT----- 442
 QY 457 TCGATCGTTTAACTTCTCAATATCTTTTAAAGCTTCTCCCTCTTTAGTTTCTTTCTTAA 516
 Db 443 -----TAAAGATTTGAAT 495
 QY 517 ATTAACCTTAAATTAACAACCTTACATATATATATATATATATATATATATATATATATATAT 576
 Db 496 ATAACCTTGGTTAATG-----TCTATTTGCATGCAACAATAATGAGGACTACTAGG 547
 QY 577 CATATATAGCTTATGTATTTTAACTATATATATATATATATATATATATATATATATATATAT 636
 Db 548 ATTATTTATTTATTTTAACTAT 607
 QY 637 GGAGAGAAAGAGTGGTACTTCTATGTGCAAGAGATCGGAATATATAGAAATGAGATAGA 696
 Db 608 GGAGAGAAAGAGTGGTACTTCTTCCAGTCCAGAGACAGGAAGTACCCGAATGGTTTCACGA 667
 QY 697 CCGAACCGAGTAAGCACTTCAGGATATTTGGAAGCCACCGAGCTGTATAGGATGATCAGA 756
 Db 668 CCAACAGAGTTGCGGGTTCTGGGTATTTGGAAGCCACTGGAATGATGATGATGATGATGATGAT 727
 QY 757 TCGGAGACTTCTCGGCCCTTATCGGATTTAAAGAAAAACCCCTAGTTTCTTCTCTCTCTCTCT 816
 Db 728 ACTGA---AGGTAGAAAAAGTTGGCATAAAAAAGCACTTGTCTTCTCTCTCTCTCTCTCTCT 784
 QY 817 CTTAAAGCACTCTGACTAGTTGGATCATGAACGATATC 856
 Db 785 CCCAAAGGCTCCAAACCCATTGGATCATGCACGATATC 824

RESULT 14

AD062518
 ID AD062518 standard; DNA; 879 BP.

XX AC AD062518;

XX DT 15-JUL-2004 (first entry).

XX DE Transcription factor G1412 orthologous sequence, SEQ ID 985.

XX KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development; ds.

XX OS Glycine max.

XX FN WO2004031349-A2.

XX PD 15-APR-2004.

XX PF 18-SEP-2003; 2003WO-US030292.

XX XX

CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

2000

SQ Sequence 1297 BP; 385 A; 275 C; 308 G; 329 T; 0 U; 0 Other;

4.48: Query Match Score 114.2: DB 13: Length 1297:

Query Match 4.4%, SCORE 114.2, DS 13, Pengine 1237,
Best Local Similarity 66.4%: Pred. No. 5.6e-13:

best local similarity 88.4%; FREQ. NO. 3.0E-15;
 Matches 180; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 597 TTAATCATAGATATATGTATATGCAGCTATGCCGCCGATAGGAGAGAAAGAGTGGTACTT 656

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 657 CTATGTGCCAAGAGATCGGAAATATAGAAATGGAGATAGACCGAACCGACTTC 716

Db 72 TTTCAGCCCTAGAGACAGGAAATACCCCTAACGGGTCCCGACCCCAACAGAGTAGCCGGTC 131

Db 132 GGGTTATTGGAAAGCCACCGGAACCGACATCATCACCACCGA--AGGTAGAAAGT 188

QY 777 CGGATTAAAGAAACCCCTAGTTTCTACTCTGGTAAAGCCCCCTAAAGGCACCTCGTACTAG 836

Search completed: April 6, 2006, 09:38:31
Job time : 1516 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2006, 09:13:28 ; Search time 9249 Seconds
(without alignments)
13182.737 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctccac.....ctaataattctttaagtaa 2606

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 segs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	466	17.9	507	9	AY201074
2	450	17.3	901	8	DR749910
3	432.8	16.6	506	1	AI993449
4	395.8	15.2	690	9	BH481683
5	333	12.8	838	9	BH666975
6	323.2	12.4	622	1	AV827617
7	314.2	12.1	947	8	DR749909
8	292.8	11.2	427	1	AV440745
9	292.8	11.2	468	1	AV442690
10	229	8.8	283	8	R84139
11	226.2	8.7	729	9	BZ028065
12	219.2	8.4	702	9	BH999674
13	214.4	8.2	280	9	AY201073
14	202.4	7.8	720	9	BZ028439
15	185.4	7.1	869	10	CZ981915
16	184.4	7.1	516	3	BI893633
17	183.2	7.0	586	9	CC955522
18	180	6.9	605	10	AG223605
19	178.2	6.8	623	8	CX527114
20	178.2	6.8	648	1	AW683672
21	178.2	6.8	653	2	BE204243
22	176.8	6.8	741	7	CO096386

23	172.4	6.6	813	6	CD575840
24	171.4	6.6	525	8	DT006901
c	171	6.6	718	10	CM220628
26	169.8	6.5	174	11	CR399335
c	167.6	6.4	813	9	AV126200
28	166.4	6.4	310	1	AV428075
29	164	6.3	1404	10	CL960332
30	163.6	6.3	853	8	DR928151
31	163.6	6.3	906	8	DR935181
32	162	6.2	880	8	DR938316
33	161	6.2	555	1	AW443468
34	153.6	5.9	486	8	DT004645
c	153.2	5.9	508	10	CG269114
36	153.2	5.9	621	10	CG136476
c	153.2	5.9	759	10	CG192989
38	153.2	5.9	822	10	CG178208
c	153.2	5.9	823	9	CC659605
c	153.2	5.9	894	10	CG178210
c	152.8	5.9	657	2	BI268158
c	151	5.8	431	1	AV798788
43	149.2	5.7	649	7	CK077495
44	149.2	5.7	770	6	CB621762
45	149.2	5.7	784	6	CB621699

ALIGNMENTS

RESULT 1
AY201074

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AY201074 507 bp DNA linear GSS 03-JAN-2005
Arabidopsis thaliana Landsberg DNA Arabidopsis thaliana
genomic clone GT109.D5.08.02.99.b.507, genomic survey sequence.
AY201074
AY201074.1 GI:27897028
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 507)
May B.P., Simorowski J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
Unpublished (2004)
Contact: Martienssen RA
Cold Spring Harbor Laboratory
1 Bungtown Rd., Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8322
Fax: 516 367 8369
Email: martiens@cshl.org
Arabidopsis thaliana sequence flanking Ds end of Ds-Trap insertion
from line GT109.
Class: transposon-tagged.
Location/Qualifiers
1..507
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="ecotype Landsberg"
/db_xref="taxon:3702"
/clone="GT109.D5.08.02.99.b.507"
/clone_lib="Arabidopsis thaliana Landsberg DNA"

ORIGIN

Query Match 17.9%; Score 466; DB 9; Length 507;
Best Local Similarity 98.4%; Pred. No. 1.3e-81;
Matches 492; Conservative 0; Mismatches 5; Indels 3; Gaps 2;
QY 457 TCGATCGTTAACAATTCCTCAATATCTTTAAACGCTTCTCCCTTTAGTTCTTTCTTAA 516
|||||
DB 10 TCGATCGTTAACAATTCCTCAATATCTTTAAACGCTTCTCCCTTTAGTTCTTTCTTAA 69
|||||

QY	517	ATTAAACCTAAATTTAAACAAACCTACATATATATCATAAAGATATACAAATATGTGTGATG- TTT	575
Db	70	ATTAAACCTAAATTTAAACAAACCTACATATATATCATAAAGATATACAAATATGTGTGATGTTTT	129
QY	576	TCATAATTTAGCTTATGTATGTATTTAATCATAGATATATGTATATGCAGCTATATGCGCGCAT	635
Db	130	TCATAATTTAGCTTATGTATGTATTTAATTTATAGATATATGTATATGCAGCTATATGCGCGCAT	189
QY	636	AGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAATATAGAAATCGGAGATAG	695
Db	190	AGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAATATAGAAATCGGAGATAG	249
QY	696	ACCGAAACCGAGTAACGACTTTCAGGATATTGGAAAGCCACCGGAGCTGATAGGATGATCAG	755
Db	250	ACCGAAACCGAGTAACGACTTTCAGGATATTGGAAAGCCACCGGAGCTGATAGGATGATCAG	309
QY	756	ATCGGAGACTTCTCGGCCCTATCGGATTTAAAGAAAAACCTAGTTTTCTACTCTGGTAAAGC	815
Db	310	ATCGGAGACTTCTCGGCCCTATCGGATTTAAAGAAAAACCTAGTTTTCTACTCTGGTAAAGC	369
QY	816	CCCTAAAGSCATCTGCTACTAGTTGGATCATCAACGAGTATCGTCTCCCGCACCATCAAAAC	875
Db	370	CCCTAAAGSCATCTGCTACTAGTTGGATCATCAACGAGTATCGTCTCCCGCACCATCAAAAC	429
QY	876	CGAGAAGTACCAAAAGGTATAAATCTTACTATAAATCTTATATATATCTTATTCATACATA	935
Db	430	CGAGAAGTACCAAAAGGTATAAATCTTACTATAAATCTTATATATATCTTATTCATACGTA	487
QY	936	CATAGATATAACCCCTAGCTA	955
Db	488	CATAGATATAACCCCTAGCTA	507

RESULT 2	DR749910	901 bp	mRNA	linear	EST 19-JUL-2005
LOCUS	85-L020255-065-002-E11-SeIB MP12-ADIS-065d	Arabidopsis thaliana			
DEFINITION	cDNA clone 002-E11, mRNA sequence.				
ACCESSION	DR749910				
VERSION	DR749910.1	GI:71035250			
KEYWORDS	EST.				
SOURCE	Arabidopsis thaliana (thalé cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;				
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE	1 (bases 1 to 901)				
AUTHORS	Paz-Ares,J., Valencia,A., Costantino,P., Vittorioso,P., Davies,B., Gilmartin,P., Giraudat,J., Parcy,F., Reindl,A., Sablowski,R., Coupland,G., Martin,C., Angenent,G.C., Baumelein,H., Mock,H.P., Carbonero,P., Colombo,L., Tonelli,C., Engstroem,P., Droegge-Laser,W., Gatz,C., Kavanagh,T., Kushnir,S., Zabeau,M., Laux,T., Holdsworth,M., Ruberti,I., Ratcliff,F., Smeekens,S., Somschik,I., Weisshaar,B. and Traas,J.				
TITLE	REGIA, an EU project on functional genomics of transcription factors from Arabidopsis thaliana				
JOURNAL	Comp. Funct. Genomics 3 (2), 102-108 (2002)				
COMMENT	Contact: Paz-Ares, Costantino, Vittorioso, Davies, Gilmartin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenent, Baumelein, Carbonero, Colombo, Tonelli, Engstroem, Droegge-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Laux, Holdsworth, Ruberti, Smeekens, Somschik, Weisshaar, Traas Bielefeld University, Institute for Genome Research Universitaetsstrasse 25, D-33594 Bielefeld, Germany Email: bernd.weisshaar@uni-bielefeld.de GI: 20202450; SeqAnalysis: undetermined frame; Translation: no full cds detected				
	Data analysis performed in the frame of REGULATORS (Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance), a Trilateral Co-Operation in Plant Genomics between Spain (MCVT), France (GENOPLANTE) and Germany (GABI) coordinated by				

G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent Thareau (JBP-Orsay UMR8618 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and Alain Leclercy (URGV-Eury UMD-IRNA-CNRS-UEVE, leclercy-ad-ibp.u-psud.fr).

Definition of the terms used to describe the quality of the clones: The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGRv5 annotation, and the resulting AGI code is presented if more than 90 percent identity was found. The sequences were also blasted against all TIGRv5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for Seqanalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGRv5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGRv5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.

Submitted data.
Insert Length: 901 Std Error: 0.00
Seq primer: Se1B GTAACATCAGAGATTTTGAGACAC.

FEATURES
source

1. .901
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:145251"
/db_xref="taxon:3702"
/clone="002-E11"
/lab_host="E. coli DH5alpha"
/clone_lbb="WPMZ-DH5-065d"
/notes="Vector: pDONR201. In the project REGIA (QLG-C71l1999-00876 Paz-Ares), a set of transcription factors generated. The ORFs were produced in the labs of the participants. Generated by RT-PCR using cDNA from tissues as a template. Initially yeast recombination to move the vectors into target constructs. Of the clones contain 'REGIA' tag: Spr-AATTCAGCTACACC-3pr; RG-tag: Spr-CATTGGCAATCCCGGATC-3pr). D. project, the GATEWAY system became all ORFs were transferred into Gateway vector. In the clones of the REGIA project, the clones DNA. For end-sequencing at the MPI of Plant Breeding Research (Weissenhofstr., the plasmids were retransformed into DH5alpha, arrayed into 96-well plates was prepared. Re-transformation because a number of clones displayed no colonies. The submission has been handled by Genebank." (http://gabi.trrdd.de).")

ORIGIN

	Query Match	17.3%;	Score 450;	DB 8;	Length 901;
	Best Local Similarity	81.3%;	Pred. NO. 1.9e-78;		
	Matches 583;	Conservative	0;	Mismatches 30;	Indels 105; Gaps 1;
Qy	1883	GTATCAAGTCTCTAAATCAATGTCTCATGTATTAATTTATAGGCTGAATAATCATCTGCCG	1942		
Db	716	GTATCGTCTTCGGTACATGAACCGAAGAGTACCAAAAGGCTGAATAATCATTTGCCG	657		
Qy	1943	AGTGTACAAAGGCCAGGAGTAGAAGATCATCATCGGTACCAAGTTCTCTCTCCACAAG	2002		

```

656 AGTGTACAAAAGGCGAGGAGTAGAAGATCATCCATCGGTACACGTTCTCTCTCCACAAG 597
2003 ACATCATAAACATTAACATCATCGATCATCCGTTTAGCCTTAAGACAAACAAACACCA 2062
596 ACATCATAAACATTAACATCATCGATCATCCGTTTAGCCTTAAGACAAACAAACACCA 537
2063 TTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2122
536 TTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 477
2123 CGAGAAGCTCTCCACCGAATATTCGCGGAGCGGAGCAGCAACAAACAAACAAACAG 2182
476 CGAGAAGCTCTCCACCGAATATTCGCGGAGCGGAGCAGCAACAAACAAACAAACAG 417
2183 TAACCTCTGAGCTTACCACTGCTCTAGCCAAATCAAAACATATATCGTCCAAATGCTTACGA 2242
416 TAACCTCTGAGCTTACCACTGCTCTAGCCAAATCAAAACATATATCGTCCAAATGCTTACGA 357
2243 CACAAGCAACAAACATGATGATGCTCTTACGAGAAATCATCAAGACGATGATGAACTGC 2302
356 CACAAGCAACAAACATGATGATGCTCTTACGAGAAATCATCAAGACGATGATGAACTGC 297
2303 CATTGTTGAGCATCTTCAAAGACTAGTTAACTACCAAAATATCAGATGGAATCAATCAA 2362
296 CATTGTTGAGCATCTTCAAAGACTAGTTAACTACCAAAATATCAGATG----- 250
2363 TCACCAATATCTTCAAATGCTCAACAGTTTCATCATACTCAACAAACAAATGCTAAGC 2422
249 ----- 250
2423 AAACGCATTACAATTGGTGGTGGCGGATGATGAGGAGCAACGCTTAATGCTCAAACTCA 2482
249 -----CAGCGACAAACGCTTAATGCTCAAACTCA 222
2483 AGCGGCGTTAGCTATGAACATGATTCCTGCGAGAACGATTCCAAACAAATGCTTGTGGGA 2542
221 AGCGGCGTTAGCTATGAACATGATTCCTGCGAGAACGATTCCAAACAAATGCTTGTGGGA 162
2543 TATGTGGAATCCAAATAGTACAGATGGAACAGAGATCACTACTACTATATTCCTTTT 2600
161 TATGTGGAATCCAAATAGTACAGATGGAACAGAGATCACTACTACTATATTCCTTTT 104

```

```

RESULT 3
A1993449
LOCUS
DEFINITION
701496358 A. thaliana, Ohio State clone set Arabidopsis thaliana
cDNA clone 701496358, mRNA sequence.
A1993449
VERSION
A1993449.1 GI:5840354
EST.

```

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 506)
Chen, J., Montoya, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D.,
Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C.,
Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shan, S.,
Nobrega, A., Murrey, J., Turner, C., Krikorian, S., Elder, L. and
Hanson, D.

Arabidopsis thaliana Gene Expression MicroArray

Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324

FEATURES

source

Email: service@genomesystems.com.
Location/Qualifiers
1..506
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="701496358"
/notes="cDNA library was made from selected clones from the
Arabidopsis thaliana Ohio State clone set."

ORIGIN

Query Match 16.6% Score 432.8; DB 1; Length 506;
Best Local Similarity 90.6%; Pred. No. 4.8e-75;
Matches 464; Conservative 0; Mismatches 41; Indels 7; Gaps 2;
QY 1992 CTCTCCCAAGACATCATTAACCATTAATCATCGACATCATCCCGTTTAGCCTTAAGCAA 2051
Db 1 CTCTCCCAAGACATCATTAACCATTAATCATCGACATCATCCCGTTTAGCCTTAAGCAA 54
QY 2052 CAACAACACCATTCATCT 2111
Db 55 CAACAACACCATTCATCT 114
QY 2112 ATCAACAATCTCGAGAAGCTCTCCACCGAATATTCGCGGAGCGGAGCAGCAACAAACAG 2171
Db 115 NNN 174
QY 2172 ACCAACAACAGTAACTCTGAGCTTACATGCTCTAGCCAAATCAAAACATATATCGTCCA 2231
Db 175 ACCAACAACAGTAACTCTGAGCTTACATGCTCTAGCCAAATCAAAACATATATCGTCCA 234
QY 2232 ATGCTTTAGCAGCAACAGCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 2291
Db 235 ATGCTTTAGCAGCAACAGCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 294
QY 2292 GATGAACATGCCAATGTTGACGATCTTCAAAGACTAGTTAACTACCAAAATATCAGATGGA 2351
Db 295 GATGAACATGCCAATGTTGACGATCTTCAAAGACTAGTTAACTACCAAAATATCAGATGGA 354
QY 2352 GGTAAACATCAATCAACCAATCTTCAAAATGCTCAACAGTTTCATCATCTACTCAACAACA 2411
Db 355 GGTAAACATCAATCAACCAATCTTCAAAATGCTCAACAGTTTCATCATCTACTCAACAACA 414
QY 2412 ATGCTTAAACGCAACGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2471
Db 415 ATGCTTAAACGCAACGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
QY 2472 CCTCAAACTCAA-GCGGCGTTAGCTATGAACA 2502
Db 475 CCTCAAACTCAA-GCGGCGTTAGCTATGAACA 506

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

BH481683
BOGNE16TR BOGN Brassica oleracea genomic clone BOGNE16, genomic
survey sequence.
BH481683
BH481683.1 GI:17689787
GSS.
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 690)
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,
Utterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490

COMMENT Other GSSs: BOGNE16TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
1. .690
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGNE16"
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match 15.2%; Score 395.8; DB 9; Length 690;
Best Local Similarity 80.3%; Pred. No. 1e-67;
Matches 561; Conservative 0; Mismatches 107; Indels 31; Gaps 7;
QY 1875 TGATATATGATCAAGTCTCTTAATCAATGCTGCTCATGTATA-ATTATAGGCTGAATATC 1933
Db |||||
QY 19 TAATATATATGATCAAGTCTCTGATCAATGTTCTATATATATATATAGGCTGAATATC 78
Db |||||
QY 1934 ATTGTGCGGAGTGTCACAAAGCCAGGAGTGAAGATCATCCATCGGTACCAAGTCTCT 1993
Db |||||
QY 79 ATTGTGCGGAGTGTCACAAAGCCAGGAGTGAAGATCATCTATCTTACCAGTCTCT 138
Db |||||
QY 1994 CTCACAAGACATCAATCAATCAATCATCGACATCATCCGCTTTAGCCTTAAGACAAC 2053
Db |||||
QY 139 ATCCACAGACATCAATCAATCAATCATCTTCCGCTTTGGCTTAAGACAACAA 198
Db |||||
QY 2054 ACAACAGCATTCCTCTTAATCAATTCGACAAACACCTTTACACAAACAAAT 2113
Db |||||
QY 199 ACAACACCACTCATCATCAATCAATTCAGACAAACACCTTAACAA 249
Db |||||
QY 2114 CAACAATCTCGAGAGCTCTCCACCGAATATTCGCGGACGCGACACAAACACGAC 2173
Db |||||
QY 250 CAACAATCTTGAGAGCTCTCAACCGAATATTTGGCGACGCTAGC---AGCATACAC 306
Db |||||
QY 2174 CAACAAGTAACTCTGAGCTTACCAATTCCTTAGCCAAATCAAAATATATCGTCCAA 2233
Db |||||
QY 307 CAACAAGTAACTCTGAGCTCACTATAGCTTTAGCCAAATCAAAATATATCGTCCAA 366
Db |||||
QY 2234 GCCTTACGACAAAGCAACAAACATTTGATGCTCTACGAGAAATCATCAAGACGATGA 2293
Db |||||
QY 367 GCCTTATGTTGTAAGCAACACCAATGAT-----CTCCACAATCAATTAAGAGATGA 420
Db |||||
QY 2294 TGAAGTGCATTTGAGAGTCTCAAGAGTGTAACTACTCAAAATATCATAGATGGAG- 2352
Db |||||
QY 421 TGAAGTGCATTTGAGAGTCTTCAAGAGTCTGTTAACTACCAAAATATCATGTTGGAG 480
Db |||||
QY 2353 --GTAACATCAATCAACCAATCTTTCAAAATTTGCTCAACAGTTTCTATCTATCAACA 2410
Db |||||
QY 481 TAGTAACATCAATCAACCAATCTTCAAAATTTGCTCAACAGTTTCTATAT- 534
Db |||||
QY 2411 AAATGCTAACGCAAGCAATTAATTTGTTGGTGGTGGCGGAGCTACAGCGCAACCGTAA 2470
Db |||||
QY 535 AGAGGTAAATGCAACCGCTTGAATTTGGTGGCGGGGCAACTACAGTG---GCGCTAAC 591
Db |||||
QY 2471 GCCTCAACTCAAGCGGCTTACCTATGACATGATTTCTGCGAGGACGATTTCCAAACAA 2530
Db |||||
QY 592 GCCTCAACCGCGGGGCAATAGCGATGAAATGATGATCCCTGCGAGGACGATTTCCAAAC 651
Db |||||
QY 2531 TCGTTTGTGGGATATGTGGAATCCAAATAGTACGAGATGG 2569
Db |||||
QY 652 TGGCTTGTGGGATCTATGGAATCACTAGTACCGGATGG 690

RESULT 5
BH666975
LOCUS
DEFINITION
BOMKF40TR BO 2.3 KB Brassica oleracea genomic clone BOMKF40, linear GSS 19-FEB-2002
genomic survey sequence.
ACCESSION
BH666975
VERSION
BH666975.1 GI:18726179
SOURCE
GSS.
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
AUTHORS
1 (bases 1 to 838)
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,
Utterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.
TITLE
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
JOURNAL
Genome Res. 15 (4), 487-495 (2005)
PUBMED
15805490
COMMENT
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
1. .838
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOMKF40"
/clone_lib="BO 2.3 KB"
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match 12.8%; Score 333; DB 9; Length 838;
Best Local Similarity 80.1%; Pred. No. 2.7e-55;
Matches 507; Conservative 0; Mismatches 80; Indels 46; Gaps 8;
QY 1983 CCAGTTCCTCTCTCCACAGACATCAATACCAATCACTCATCATCATCATCATCATCATCAT 2042
Db |||||
QY 2043 TTAAAG-----ACAAACAACACACCATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2096
Db |||||
QY 62 GAATGACAAACAACAACAACAACACCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121
Db |||||
QY 2097 AACACAACAACAACAACAACAACAATCTCGAGAGTCTCCACGAATATTTCCGGCGACGCG 2156
Db |||||
QY 122 AACAA-----CAACAATCTCGAGAGTCTCTCAACGAATATTTCCGGCGACGCT 169
Db |||||
QY 2157 AGCAACAACAACGACCAACAACAAGTAACTCTGACGTTTACCATTGCTCTAGCCAAATCAA 2216
Db |||||
QY 170 AGC---ACCATAACCACTACAAACAGTAATTTCTGACGTCACCATAGCTTTAGCCAAAC 226
Db |||||
QY 2217 AACATATATCTCAATGCCCTTACGACAAACAACAACAATCATCATCATCATCATCATCAT 2276
Db |||||
QY 227 AACATCTATCTCAATGCCCTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 274
Db |||||
QY 2277 AATCATCAAGACGATGATGAAGTGCCTTCTGACGATCTTCAAGACTAGTTAACTA- 2335
Db |||||
QY 275 TCTAATCAAGAGATGATGAAGACCGCAATAGTTGATGATCTTCAAGACTCTGTTAACTAC 334
Db |||||
QY 2336 --CCAATATCAGATGGAGGTAAACATCAATCAACCAATATCTTCAATTTGCTCAACAGTTT 2393
Db |||||

Db 335 GGCACAAATATCTGGTGGAGTAACATCAATCAACCAATCACTAAATGGCTCAACAGTTT 394
 QY 2394 CATCATCTACACACAAATGCTTAACGCAACGCAATACATTTGGTGGCTGGCGGACT 2453
 Db 395 CATATCAACAACTACTA-----AACGCAATGCGTTGCAATCGGTGACGCGCAGACT 448
 QY 2454 ACAGCGCAACGGCTAATGCTCAAACTCAAGCGGCGTTAGCTATGAAATGATTCCTGCA 2513
 Db 449 ACAG--CAGCGTTAACGCGCTCAACAGCGGCGACGTTAGCGATGAACATGATCCCTGCA 505
 QY 2514 GGAACGATTCMAACAATGCTTTGTGGGATATGTGGAATCCAAATAGTACCGAGTGAAC 2573
 Db 506 GGGACGATTCMAACAATGCTTTGTGGGATATGTGGAATCCAAATAGTACCGAGTGAAC 564
 QY 2574 AGAGATCACTATCAATATTCCTTTTAAAGTAA 2606
 Db 565 AAGATCATTTACTAATATTCCTTTTAGATGAA 597

RESULT 6
 LOCUS AV827617 622 bp mRNA linear EST 01-APR-2002
 DEFINITION AV827617 RAF19 Arabidopsis thaliana cDNA clone RAF109-17-121 5', mRNA sequence.

ACCESSION AV827617 GI:19869677

VERSION AV827617

KEYWORDS EST

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 622)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.

Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished (2002)

CONTACT: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060

Email: msekic@r.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES
 source Location/Qualifiers
 1..622

/organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF109-17-121"

/dev stages="plants at various developmental stages from germination to mature seeds"
 /lab host="DH10B"
 /clone_lib="RAF19"

/note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

ORIGIN
 Query Match 12.4%; Score 323.2; DB 1; Length 622;
 Best local Similarity 93.6%; Pred. NO. 2.3e-53;
 Matches 337; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1883 GTATCAAGTCTCTAATCAATGCTCATGTATATATAGGCTGAATATCAATGCTGCGC 1942

Db 263 GTATCGTCTTCGACCATGAACCGAGAAGTACCAAGGCTGAATATCATTTGTCGC 322
 QY 1943 AGTGTACAAAAGGCGAGAGTAGAAGATCATCCATCGTACACAGTTCTCTCTCCACAAG 2002
 Db 323 AGTGTACAAAAGGCGAGAGTAGAAGATCATCCATCGTACACAGTTCTCTCTCCACAAG 382
 QY 2003 ACATCATAACCATTAACATCATCGACATCATCCGTTTAGCCTTTAGACAAACACACACCA 2062
 Db 383 ACATCATAACCATTAACATCATCGACATCATCCGTTTAGCCTTTAGACAAACACACACCA 442
 QY 2063 TTATCTCTCTCTTAATCATTCGCAACAACTTAACAAACACACACATCAACAATCT 2122
 Db 443 TTATCTCTCTCTTAATCATTCGCAACAACTTAACAAACACACATCAACAATCT 502
 QY 2123 CGAGAAGCTCTCCACCGAATATTCGGCGAGCGGAGACACACACACACACACACAG 2182
 Db 503 CGAGAAGCTCTCCACCGAATATTCGGCGAGCGGAGACACACACACACACACACAG 562
 QY 2183 TAACTCTGACGTTACCATTTGCTTAGCCAAATCAAAACATATATGCTCAATGCTTAGCA 2242
 Db 563 TAACTCTGACGTTACCATTTGCTTAGCCAAATCAAAACATATATGCTCAATGCTTAGCA 622

RESULT 7

LOCUS DR749909

DEFINITION 85-L020254-065-002-E11-Sela MP1Z-ADIS-065d Arabidopsis thaliana

CDNA clone 002-E11, mRNA sequence.

ACCESSION DR749909

VERSION DR749909.1

KEYWORDS EST

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 947)

Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B., Gilmartin, P., Giraudat, J., Parcy, F., Reindl, A., Sablowski, R., Coupland, G., Martin, C., Angenot, G.C., Baeumlein, H., Mock, H.P., Carbonero, P., Colombo, L., Tonelli, C., Engstrom, P., Droegge-Laser, W., Gatz, C., Kavanagh, T., Kushnir, S., Zabeau, M., Laux, T., Hordsworth, M., Ruberti, I., Ratcliff, F., Smeekens, S., Somssich, I., Weisshaar, B. and Traas, J.

REGIA, an EU project on functional genomics of transcription factors from Arabidopsis thaliana

Comp. Funct. Genomics 3 (2), 102-108 (2002)

Contact:

Paz-Ares, Costantino, Vittorioso, Davies, Gilmartin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenot, Baeumlein, Carbonero, Colombo, Tonelli, Engstrom, Droegge-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Laux, Hordsworth, Ruberti, Smeekens, Somssich, Weisshaar, Traas

Bielefeld University, Institute for Genome Research

Universitaetsstrasse 25, D-33594 Bielefeld, Germany

Email: bernd.weisshaar@uni-bielefeld.de

AGI: AT2G02450; Seqanalysis: undetermined frame; Translation: no full cds detected

Data analysis performed in the frame of REGULATORS (Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance), a Trilateral Co-Operation in Plant Genomics between Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent Thureau (IBP-Orsay UMR8618 CNRS-UPS, thureau-ad-ibp.u-psud.fr) and Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE, lecharny-ad-ibp.u-psud.fr).

Definition of the terms used to describe the quality of the clone: The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGR5 annotation, and the resulting AGI code is presented if more

than 90 percent identity was found. The sequences were also blasted against all TIGR5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for Seqanalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGR5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGR5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.

Insert Length: 947 Std Error: 0.00
Seq primer: Seta TCGGTTAAGCTAGCATGATCTC.

FEATURES source

Location/Qualifiers
1..947
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:1452250"
/db_xref="taxon:3702"
/clone="002-E11"
/lab_host="E. coli DH5alpha"
/clone_lib="MP12-ADIS-065d"
/note="Vector: pDONR201; In the context of the EU-funded Project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various A. thaliana tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG_tag1: Spr-AATTCAGCTCACACC-3pr; RG_tag2: Spr-CATGCAATCCGGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weisshaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by GabiPD (http://gabi.rzpd.de)."

ORIGIN

Query Match 12.1%; Score 314.2; DB 8; Length 947;
Best Local Similarity 92.2%; Pred. No. 1.4e-51;
Matches 331; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1883 GTATCAAGTCTCTTAATCAATGTCTCATGTATATATATAGGCTGAATATCATTTGGCGG 1942
DB 589 GTATCGTCTTCGTACCATGMAACCGAGAAGTACCAAAAGGCTGAAATATCATTTGGCGG 648

QY 1943 AGTGTAACAAAGGCCAGGAGTAGAAGATCATTCATCGGTACCAAGTTCTCTCCACAG 2002
DB 649 AGTGTAACAAAGGCCAGGAGTAGAAGATCATTCATCGGTACCAAGTTCTCTCCACAG 708

QY 2003 ACATCATACCAATAACTCATCGACATCATCCCGCTTTAGCGCTTTAAGACAACAACACCA 2062
DB 709 ACATCATACCAATAACTCATCGACATCATCCCGCTTTAGCGCTTTAAGACAACAACACCA 768

QY 2063 TTTCATCTCTCTTAATCAATTCGACAAACACCTTAACAAACAAACAAACAAACAAATCT 2122
DB 769 TTTCATCTCTCTTAATCAATTCGACAAACACCTTTAACAACAACAACAACAACAATCT 828

QY 2123 CGAGAAGCTCTCCACCGAATATTCGGCGAGCAGCAACAACAGCACAACAG 2182
DB 829 CGAGAAGCTCTCCACCGAATATTCGGCGAGCAGCAGCAACAACAGCACAACAG 888

QY 2183 TAACTCTGACGTTACCAATGCTGTAGCCATCAAAACATATATGTCCTCAATGCCTTACG 2241
DB 889 TAACTCTGACGTTACCAATGCTGTAGCCATCAAAACATATATGTCCTCAATGCCTTACG 947

RESULT 8
AV440745 427 bp mRNA linear EST 18-FEB-2004
LOCUS AV440745 Arabidopsis thaliana above-ground organ two to six-week old Arabidopsis thaliana cDNA clone AP208c08_f 3', mRNA sequence.
ACCESSION AV440745
VERSION AV440745.1 GI:7611116
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (Bases 1 to 427)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..427
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="AP208c08_f"
/tissue_type="aboveground organs"
/dev_stages="two to six-week old"
/clone_lib="Arabidopsis thaliana above-ground organ two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 11.2%; Score 292.8; DB 1; Length 427;
Best Local Similarity 99.3%; Pred. No. 2.3e-47;
Matches 294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCAATTGTATCTCTCCACCAAGCATCATTCCTCATGATTAACCAAGTCAACATAAC 60
DB 32 ATGGCAATTGTATCTCTCCACCAAGCATCATTCCTCATGATTAACCAAGTCAACATAAC 91

QY 61 GAAAAGGTATAGAGACAATGATCATAGAGCGCGCAAGAGAGTCATGTCCCAATGAA 120
DB 92 GAAAAGGTATAGAGACAATGATCATAGAGCGCGCAAGAGAGTCATGTCCCAATGAA 151

QY 121 GATGAAGCTGATGATCATGATCATGATGTCATGTCCTCCCGGATTTAGATTCCTCTACC 180
DB 152 GATGAAGCTGATGATCATGATCATGATGTCATGTCATGTCCTCCCGGATTTAGATTCCTCTACC 211

QY 181 GAAGAAGAACTCATAGAGTTTAACTTCCTCCGCAAAAGTTGAAGGCAACGCTTTTATGTA 240
DB 212 GAAGAAGAACTCATAGAGTTTAACTTCCTCCGCAAAAGTTGAAGGCAACGCTTTTATGTA 271

QY 241 GAATCATCATCTTCTCTCGATCTTTATCGCTATGATCTCTTGGAACTTCTCTGTA 296
DB 272 GAATCATCATCTTCTCTCGATCTTTATCGCTATGATCTCTTGGAACTTCTCTGTA 327

RESULT 9
 AV442690/c
 LOCUS
 DEFINITION AV442690 Arabidopsis thaliana above-ground organ two to six-week old Arabidopsis thaliana cDNA clone AP208c08_r 5', mRNA sequence.
 ACCESSION AV442690
 VERSION AV442690.1 GI:7613105
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 REFERENCE 1 (bases 1 to 468)
 AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
 JOURNAL DNA Res. 7 (3), 175-180 (2000)
 PUBMED 10907847
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
 source
 1..468
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="AP208c08_r"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /clone_lib="Arabidopsis thaliana above-ground organ two to six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
 Query Match 11.2%; Score 292.8; DB 1; Length 468;
 Best Local Similarity 99.3%; Pred. No. 2.3e-47;
 Matches 294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 ATGCAATTGATCTCTCCACCAACAGCATCATCTCCATGATGTAACCAAGTCAACAAATAC 60
 Db 396 ATGCAATTGATCTCTCCACCAACAGCATCATCTCCATGATGTAACCAAGTCAACAAATAC 337
 Qy 61 GAAAAGGTATAGAGCAATGATCATAGAGCGCGCCAGAGAGTCAATGTCCTCAAAATGAA 120
 Db 336 GAAAAGGTATAGAGCAATGATCATAGAGCGCGCCAGAGAGTCAATGTCCTCAAAATGAA 277
 Qy 121 GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 Db 276 GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 217
 Qy 181 GAAGAAGCACTCATAGAGTTTTTACCTTCGCCGAAAAGTTGAAGCAACGCTTTAATGTA 240
 Db 216 GAAGAAGCACTCATAGAGTTTTTACCTTCGCCGAAAAGTTGAAGCAACGCTTTAATGTA 157
 Qy 241 GAATCATCACTTCTCGATCTTTATCGCTATGATCTTGGGAACTTCTGGTAA 296
 Db 156 GAATCATCACTTCTCGATCTTTATCGCTATGATCTTGGGAACTTCTGGTAA 101

RESULT 10
 R84139
 LOCUS
 DEFINITION 16098 Lambda-PRL2 Arabidopsis thaliana cDNA clone 154014T7, mRNA sequence.
 ACCESSION R84139

VERSION
 KEYWORDS
 SOURCE

ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana

REFERENCE
 AUTHORS

Newman, T., deBruijn, F. J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E., and Somerville, C.
 Genes Galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 7846151

TITLE

JOURNAL
 PUBMED

COMMENT

On Jan 9, 1998 this sequence version replaced gi:930573.

Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@ibm.cl.msu.edu
 Seq primer: T7 dye primer.

FEATURES
 source

1..283
 /organism="Arabidopsis thaliana"

/mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="154014T7"
 /clone_lib="Lambda-PRL2"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

ORIGIN

Query Match 8.8%; Score 229; DB 8; Length 283;
 Best Local Similarity 90.2%; Pred. No. 9.4e-35;
 Matches 249; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
 Qy 2035 GTTTAGCTTTAAGACAAACAAACACCATTCATCTCTCTTAATCATTCGACCAACACC 2094
 Db 1 GTTTAGCTTTAAGACAAACAAACACCATTCATCTCTCTTAATCATTCGACCAACACC 60
 Qy 2095 TTAACAAACAAACAAACATCAACATCTCGAGAAGCTCTCCACCGAATATTCGGCGACG 2154
 Db 61 TTAACAAACAAACATCAACAAATNTNGAGAAGCTCTCCACCGAATATTCGGCGACG 120
 Qy 2155 GCAGCAACAAACACGACCAACAAACAGTAACTCTGACGTTACCATTCCTTAGCCATC 2214
 Db 121 GCAGCAACAAACACGACCAACAAACAGTAACTCTGACGTTACCATTCCTTAGCCATC 180
 Qy 2215 AAAACATATATCTGCCAATGCTTACGACACAGCAACACATTCATAGTCTCTACGA 2274
 Db 181 AAAACATATNTGGTCCCAATGCTTATGGCAGCAACAAACATTCATAGTCTCTACGA 240
 Qy 2275 GAAATCATCAAGA-CGATGATGAACATGCCATTGTT 2309
 Db 241 GTAATCATCAAGTGGTGGTGGAGACTGCCATTTTT 276

RESULT 11
 BZ028065
 LOCUS

BZ028065 729 bp DNA linear GSS 08-OCT-2002

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2006, 09:19:16 ; Search time 457 Seconds
(without alignments)
10136.380 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctccac.....ctaatattctttaagtaa 2606

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97.4	3.7	1216	3	US-09-533-029-121
2	90.8	3.5	1212	3	US-09-533-029-3
3	90.6	3.5	1409	3	US-09-533-029-37
4	88.4	3.4	914	3	US-09-533-029-23
5	87.8	3.4	2680	3	US-09-614-408-7
6	86.4	3.3	1211	3	US-09-614-408-1
7	86.4	3.3	1211	3	US-09-614-981-1
8	84.8	3.3	1209	3	US-09-614-408-6
9	83.8	3.2	187169	3	US-09-949-016-12776
10	83.8	3.2	191569	3	US-09-949-016-15940
11	81.4	3.1	1287	3	US-09-889-926-1
12	80.2	3.1	30820	3	US-09-949-016-17145
13	79.8	3.1	1141	3	US-09-806-708B-22
14	79.6	3.1	18773	3	US-09-949-016-14164
15	77.4	3.0	119153	3	US-09-949-016-12378
16	76.2	2.9	134987	3	US-09-949-016-15348
17	76.2	2.9	134987	3	US-09-949-016-15349
18	76.2	2.9	134987	3	US-09-949-016-15350
19	76.2	2.9	134987	3	US-09-949-016-15507
20	76.2	2.9	134987	3	US-09-949-016-15508
21	76.2	2.9	134987	3	US-09-949-016-15509
22	76	2.9	913	3	US-09-533-029-73
23	74.2	2.8	55886	3	US-09-949-016-15129
24	73.6	2.8	114139	3	US-09-949-016-16536

C 25	72.2	2.8	205044	3	US-09-949-016-15851	Sequence 15851, A
C 26	72.2	2.8	205044	3	US-09-949-016-15852	Sequence 15852, A
C 27	72.2	2.8	205044	3	US-09-949-016-15853	Sequence 15853, A
C 28	72.2	2.8	223471	3	US-09-949-016-12387	Sequence 12387, A
C 29	72.2	2.8	223471	3	US-09-949-016-12724	Sequence 12724, A
C 30	72.2	2.8	223471	3	US-09-949-016-12725	Sequence 12725, A
C 31	71.4	2.7	700	3	US-09-735-271-1038	Sequence 1038, Ap
C 32	71.4	2.7	95255	3	US-09-949-016-17067	Sequence 17067, A
C 33	71.4	2.7	231129	3	US-09-949-016-16110	Sequence 16110, A
C 34	71.4	2.7	26293	3	US-09-949-016-11934	Sequence 11934, A
C 35	71	2.7	18773	3	US-09-949-016-14164	Sequence 14164, A
C 36	70.8	2.7	1141	3	US-09-806-708B-22	Sequence 22, Appl
C 37	70.2	2.7	67755	3	US-09-949-016-13703	Sequence 13703, A
C 38	70	2.7	251672	3	US-09-949-016-17296	Sequence 17296, A
C 39	70	2.7	251682	3	US-09-949-016-11973	Sequence 11973, A
C 40	69.4	2.7	67755	3	US-09-949-016-13703	Sequence 13703, A
C 41	69.2	2.7	25590	3	US-09-949-002-777	Sequence 777, App
C 42	69	2.6	462589	3	US-09-949-016-12900	Sequence 12900, A
C 43	69	2.6	476044	3	US-09-949-016-12412	Sequence 12412, A
C 44	68.6	2.6	263693	3	US-09-949-016-12386	Sequence 12386, A
C 45	68.6	2.6	263694	3	US-09-949-016-16915	Sequence 16915, A

ALIGNMENTS

RESULT 1

US-09-533-029-121
; Sequence 121, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Onaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G501
US-09-533-029-121

Query Match	3.7%	Score 97.4;	DB 3;	Length 1216;
Best Local Similarity	64.2%	Pred. No. 1.9e-11;		
Matches 165;	Conservative	0;	Mismatches 86;	Indels 6;
Gaps	1;			
QY	618	TGCAGCTATGGCGCGATAGGAGAGAAAGTGGTACTTCTATGTGCAAGAGATCGGAA	677	
Db	217	TCCAGAGATGTCTCTACGGAGAGAAAGTGGTACTTCTCTACCTAGATCGGAA	276	
QY	678	ATATGAAATGGAGATAGACCGAACCGAGTAACACTTCAGGATATTTGGAAGCCCGG	737	
Db	277	ATACCCAAACGGTTCGCTCTTAACCGGCGAGCAGGAAACCGGTTATTGGAAGTACCGG	336	
QY	738	AGCTGATAGGATGATCAGATCGGAGACTTTCGGCTATCGGATTAAGAAACCCCTAGT	797	

Db 337 AGCAGATAAACCGAT-----TGGTAAACCGAAGAGCTTGGGTATCAAGAAAGCACTCGT 390
QY 798 TTTCTACTCTGTAAGCCCTTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857
Db 391 CTTCTACGAGGAAAGCTTCAAAAGGGATTAAAGACCAATTGGATAATGATGATGATGCG 450
QY 858 TCTTCGCGACCATGAAA 874
Db 451 TCTCGCTAATGTTGATA 467

RESULT 2

US-09-533-029-3

; Sequence 3, Application US/09533029

; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddle, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Marsha

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Reuber, Lynne

; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES

; FILE REFERENCE: MBI-010

; CURRENT APPLICATION NUMBER: US/09/533,029

; CURRENT FILING DATE: 2000-03-22

; EARLIER APPLICATION NUMBER: 60/125,814

; EARLIER FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 1212

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G759

; US-09-533-029-3

Query Match 3.5%; Score 90.8; DB 3; Length 1212;
Best Local Similarity 61.8%; Pred. No. 4.9e-10;
Matches 162; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

QY 599 AATCATAGATATATGTATATGAGCTATGCGCGCGATAGAGAGAAAGTGGTACTTCT 658
Db 165 AATTGCGATCCATGGGTTTACCAAAATAAGCAATTATTGAGAGAAAAGAAATGGTATTTT 224
QY 659 ATGTGCAAGAGATCGGAATATAGAAATGGAGATAGACCGAAGCAAGTAACGACTTCAG 718
Db 225 TTAGTCTTAGGATAGAAATATCCAAACGGTCAAGACTAACCGGTTGCGGATCGG 284
QY 719 GATATTGGAAGCCACCGAGCTGTATGATGATCATGATCGGAGACTTCTCGGCGCTATG 778
Db 285 GTTATTGGAAGCTACGGGTACGGATAAAATAATC---TCGACGGAAGCAAAAGATTG 341
QY 779 GATTAAAGAAACCCCTAGTTTCTACTCTGGTAAAGCCCTTAAAGCACTCTGACTAGTT 838
Db 342 GTATTAAAGAAAGCTTGGTGTGTTTACATCGGAAAAGCTCTTAAAGGTACTATAAACCAATT 401
QY 839 GGATCATGAACGAGTATCGTCT 860
Db 402 GGATCATGATGATATCGTCT 423

RESULT 3

US-09-533-029-37

; Sequence 37, Application US/09533029
; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddle, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Marsha

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Reuber, Lynne

; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES

; FILE REFERENCE: MBI-010

; CURRENT APPLICATION NUMBER: US/09/533,029

; CURRENT FILING DATE: 2000-03-22

; EARLIER APPLICATION NUMBER: 60/125,814

; EARLIER FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 37

; LENGTH: 1409

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G502

; US-09-533-029-37

Query Match 3.5%; Score 90.6; DB 3; Length 1409;
Best Local Similarity 62.7%; Pred. No. 5.6e-10;
Matches 160; Conservative 0; Mismatches 89; Indels 6; Gaps 1;

QY 618 TGCAGCTATGCGCGCGATAGAGAGAAAGTGGTACTTCTATGTGCCAAGATCGGAA 677
Db 379 TCCTGGTTTACCTTGTATGTTGAGAGGAATGGTACTTCTCTCCAGGACAGAAA 438
QY 678 ATATAGAAATGGAGATAGACCGAAGCGAGTAACGACTTCAGGATATTCGAAAGCCACCG 737
Db 439 ATATCCCAACGGTTGCGGTCCTTAACCGGTCGCTGGTTCTGTTACTGGAAGCTACCG 498
QY 738 AGCTGATAGGATGATCAGATCGGAGACTTCTCGGCGCTATCGGATTAAGAAACCCCTAGT 797
Db 499 AGCTGATAAACGATCGGA-----CTACCTAAACCGGTCGGAATTAAGAAAGCTCTTGT 552
QY 798 TTTCTACTCTGGTAAAGCCCTTAAAGGCATCTCGTACTAGTTGGATCATGAACGAGTATCG 857
Db 553 TTTCTACGCGCGCAAGCTCCAAAGGGAGAGAGAAAACCAATTGGATCATGACGAGTACCG 612
QY 858 TCTTCCGCGACCATGA 872
Db 613 TCTCGCGCGACGTTGA 627

RESULT 4

US-09-533-029-23

; Sequence 23, Application US/09533029

; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddle, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Marsha

QY	737	GAGCTGATA---GGATGATCAGATCGGAGACTTCTCGGCCCTATCGGATTAAGAAGAAACCC	793
Db	963	GAACCGACAAAGCGGTTTTTACTTCCGGTGGAAACAAAGAGTTGGGGTAAAAAGGCGC	1021
QY	794	TAGTTTTCTACTCTGGTAAAGCCCTAAAGGCACCTCGTACTAGTTGGATCATGAACGAGT	853
Db	1023	TCGTTTTTTACGGCGGTAAACCAACAAAGGGTAAAAACTAAATTGGATTCGTGCATGAAT	1081
QY	854	ATCGTCTTCGCGACCATGAAACCGAGAAGTACCAA	888
Db	1083	ACAGAGTTGTAGAAAATTAACCAAAATTAACAAAGCCA	1117

```

RESULT 6
US-09-614-408-1
; Sequence 1, Application US/09614408
; Patent No. 6762347
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: VREBALOV, JULIA
; APPLICANT: NOENSIE, FREDERICK
; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
; FILE REFERENCE: TAMK:213
; CURRENT APPLICATION NUMBER: US/09/614,408
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,357
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1211
; TYPE: DNA
; ORGANISM: Tomato
US-09-614-408-1

```

Query Match	3.3%	Score 86.4	DB 3	Length 1211
Best Local Similarity	59.9%	Pred. No. 4.3e-09		
Matches 163	Conservative	0	Mismatches 106	Indels 3
Gaps	1			
620	QY	CAGCTATGCGGCGATAGAGAGAGAAAGTGGTACTTCTATGTGCCAAGAGATCGGAAT	679	
220	DB	CTGCTAAGCAATATTCGGAGAGCAAGATGGTCTCTTTTAGTCCAAGAGATAGAAAAT	279	
680	QY	ATAGAAATCGAGATAGACCGAACCGAGTACGACTTCAGGATATTGGAAAGCCACCGGAG	739	
280	DB	ATCCTAACGGGCGAGGCCAAATCGGGCTGCAACATCGGGTTATTGGAAGGCTACCGGA	339	
740	QY	CTGATA--GGATGATCAGATCGGAGACTTCTCGGCCATCGGATTAAAGAAAAACCCCTAG	796	
340	DB	CGGACAAAGCCGGTTTTTACTTCCGGTGGAAACACAAAGAGTTGGGGTAAAAAAGGCGCTCG	399	
797	QY	TTTTCTACTCTGTAAAGCCCTAAAGGCACCTCGTACTAGTTGGATCATGAACGAGTATC	856	
400	DB	TTTTTTTACGGCGGTAAACACCAAAAGGGGTAAAACTAATTGGATCATGCTGAATACA	459	
857	QY	GTCTTCGCCACCATGAAACCGAGAAGTACCAG	888	
460	DB	GAGTCTGAGAAATAAAACAAATAAACCAAGCCA	491	

RESULT 7
US-09-614-981-1
; Sequence 1, Application US/09614981
; Patent No. 6787687
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: PADMANABHAN, VEERARAGAVAN
; APPLICANT: RUEZINSKI, DIANE
; APPLICANT: VREBALOV, JULIA
; APPLICANT: WHITE, RUTH

```

; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 914
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G503
US-09-533-029-23

```

Query Match	3.4%;	Score 88.4;	DB 3;	Length 914;
Best Local Similarity	64.1%;	Prod. No. 1.5e-09;		
Matches 150;	Conservative 0;	Mismatches 81;	Indels 3;	Gaps 1;
QY	629	CGCGGATAGGAGAGAAAGAGTGGTACTTCATGTGCCAAGAGATCGGAATATAGAAATG	688	
DB	252	CAGAGTTTGGAGAAATGAGTGGTATTCTTCAGCCCTAGAGAAAGAAATATCCAAACG	311	
QY	689	GAGATAGACCGAACCGGAGTAAACGACTTTCAGGATATTGCGAAAGCCACCGGAGCTGATAGGA	748	
DB	312	GAGTCAGACCAAAACCGGGCAGCTGTTTTCGGGTATTGCGAAAGCAACCGGTACAGACAAAG	371	
QY	749	TGATCAGATCGGAGACTTCTCGGCCATCGGATTAAAGAGAAAACCGTAGTTTTCTACTCTG	808	
DB	372	CAATTACACGC---GTTTCAGTAAACGTAGGTGTCAAGAAAGCTCTCGTCTTCTACAAAG	428	
QY	809	GTAAAGCCCTAAAGGCACCTCGTACTAGTTGGATCATGAACGAGTAGTCGCTTTC	862	
DB	429	GTAGACCTCTTAAAGGAATCAAAACCTGACTGGATCATGTCATGATAGTATCGTCTCC	482	

```

RESULT 5
US-09-614-408-7
; Sequence 7, Application US/09614408
; Patent No. 6762347
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: VREBALOV, JULIA
; APPLICANT: NOENSIE, FREDERICK
; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
; FILE REFERENCE: TAMK:213
; CURRENT APPLICATION NUMBER: US/09/614,408
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,357
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2680
; TYPE: DNA
; ORGANISM: Tomato
US-09-614-408-7

```

Query Match	3.4%	Score 87.8	DB 3	Length 2680
Best Local Similarity	60.0%	Pred. No. 2.7e-09		
Matches 165	Conservative 0	Mismatches 107	Indels 3	Gaps 1
Qy	617	ATGCAGCTATGGCGGCGATAGGAGAGAAAGATGGTGTACTTCTATGTGCCCAAGAGATCGGA	676	
Ddb	843	ACGCAGCTAAGGCGAATATTCCGAGAGCAAGATGGTCTCTTTTAGTCCAAAGAGATAGAA	902	
Qy	677	AATATAGAAATGGAGATAGACCGAACCGAGTAAACGACTTCAGAGATATTGGAAGGCCACCG	736	
Ddb	903	AATATCCTAACGGGGGAGCGCCAAATCGGGCTGCAACATCGGTATTGGAAGGCTACCG	952	

; TITLE OF INVENTION: RIN GENE COMPOSITIONS AND METHODS FOR USE THEREOF
 ; FILE REFERENCE: TAMK:214
 ; CURRENT APPLICATION NUMBER: US/09/614,981
 ; CURRENT FILING DATE: 2000-07-12
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1211
 ; TYPE: DNA
 ; ORGANISM: Tomato
 US-09-614-981-1

Query Match 3.3%; Score 86.4; DB 3; Length 1211;
 Best Local Similarity 59.9%; Pred. No. 4.3e-09;
 Matches 163; Conservative 0; Mismatches 106; Indels 3; Gaps 1;
 QY 620 CAGCTATGCGGCGATAGAGAGAAAGAGTGGTACTTCTATGTCGCAAGAGATCGGAAT 679
 DB 220 CTGCTAAGGCAATATTTCGAGAGCAAGAATGGTCTCTTTTATGTCGAAGATAGAAAAT 279
 QY 680 ATAGAATGAGATAGACCGGAGTAACCGAGTAAACGACTTCAGGATATTGGAAGCCACCGGAG 739
 DB 280 ATCTAACGGGCGAGGCCAAATCGGGCTGCAACATCGGGTTATTGGAAGGCTACCGGAA 339
 QY 740 CTGATA---GGATGATCAGATCGGAGACTTCTCGGCTATCGGATTTAAAGAAAACCCCTAG 796
 DB 340 CCGACRAGCGGTTTTACTTCGGTGGACACAAAAGGTTGGGTAAAGGCGCTCG 399
 QY 797 TTTTCTACTCTGTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATC 856
 DB 400 TTTTCTACGCGGTAAACCAACCAAGGGGTAAACAACTAATTCGATCATGCATGAATACA 459
 QY 857 GTCTCCGACCATGAACCGGAGAGTACCAA 888
 DB 460 GAGTTGTAGAAAATAAACAATAACAAGCCA 491

RESULT 8
 US-09-614-408-6
 ; Sequence 6, Application US/09614408
 ; Patent No. 6762347
 ; GENERAL INFORMATION:
 ; APPLICANT: GIOVANNONI, JAMES
 ; APPLICANT: TANKSLEY, STEVEN
 ; APPLICANT: VREBALOV, JULIA
 ; APPLICANT: NOENSIE, FREDERICK
 ; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
 ; FILE REFERENCE: TAMK:213
 ; CURRENT APPLICATION NUMBER: US/09/614,408
 ; CURRENT FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: 60/143,357
 ; PRIOR FILING DATE: 1999-07-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Tomato
 US-09-614-408-6

Query Match 3.3%; Score 84.8; DB 3; Length 1209;
 Best Local Similarity 59.6%; Pred. No. 9.5e-09;
 Matches 162; Conservative 0; Mismatches 107; Indels 3; Gaps 1;
 QY 620 CAGCTATGCGGCGATAGAGAGAAAGAGTGGTACTTCTATGTCGCAAGAGATCGGAAT 679
 DB 220 CTGGAAGGCAATATTTCGAGAGCAAGAATGGTCTCTTTTATGTCGAAGATAGAAAAT 279
 QY 680 ATAGAATGAGATAGACCGGAGTAACCGAGTAAACGACTTCAGGATATTGGAAGCCACCGGAG 739
 DB 280 ATCTAACGGGCGAGGCCAAATCGGGCTGCAACATCGGGTTATTGGAAGGCTACCGGAA 339
 QY 740 CTGATA---GGATGATCAGATCGGAGACTTCTCGGCTATCGGATTTAAAGAAAACCCCTAG 796

DB 340 CCGACRAGCGGTTTTACTTCCGTTGGAAACACAAAAGGTTGGGTAAAGGCGCTCG 399
 QY 797 TTTTCTACTCTGTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATC 856
 DB 400 TTTTCTACGCGGTAAACCAACCAAGGGGTAAACAACTAATTCGATCATGCATGAATACA 459
 QY 857 GTCTCCGACCATGAACCGGAGAGTACCAA 888
 DB 460 GAGTTGTAGAAAATAAACAATAACAAGCCA 491

RESULT 9
 US-09-949-016-12776/c
 ; Sequence 12776, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12776
 ; LENGTH: 187169
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(187169)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-12776

Query Match 3.2%; Score 83.8; DB 3; Length 187169;
 Best Local Similarity 49.6%; Pred. No. 5.7e-08;
 Matches 298; Conservative 0; Mismatches 297; Indels 6; Gaps 3;
 QY 1332 TTTAAGTTACCAACATTTTTTGGAAATAATTGGCATATATGAAATATACCAACATATT 1391
 DB 1325 TTTATTTATATATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1266
 QY 1392 ATATCGCAACATTTTAAATCTATACGAATGATACGGTTTATGGAGTAGACCGAAAAAT 1451
 DB 1265 TATATAAATATATATTTTATATATATATATAAATATATATATATATTTATATATA 1206
 QY 1452 ATTATGATACGGAATAATGCAATGAGATAGATAAATACATTTTTTTGGGCTCTTCGACT 1511
 DB 1205 TATTTATATATATATATAAATATATATATATATATATATATATATATATATATA 1146
 QY 1512 ATATGCTGTCACCATTTTGAACCAATAAATTTTATAAAATTTCTATGTATATATATAT 1571
 DB 1145 ATATATATATATATATATATATATATA-TTATATATTTTATATATATATATATATAT 1087
 QY 1572 TATGATGATGATGAAGACAGCTAAACCAACAGGGTTGACATAAATATCTATGTGTATCT 1631
 DB 1086 TATATATATATATTTTATACATATATATATATATATATATATTTTATATATATATAT 1027
 QY 1632 ATTGCACATTCCTCTGTACTAATAAATAAATAAATTTACCAATTAATATATAAATAAATA 1691
 DB 1026 TTTATATTTTATTTTAT 968
 QY 1692 TAAATATATATCTTCTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1751
 DB 967 ATATATTTTATATATATATATATATTTTATATATTTTATATTTATTTATTTATTTAT 912

Qy 1752 CTCTCAACGCTGCTATCTTTTCTGCTAACTGCTAACTTTACATAGTTTGTGCAATCTTCTT 1811
 |||||
 Db 911 TTAT 852
 |||||
 Qy 1812 TCAAAACCATATCTTCGATAAATGATATTTTTCATAGATATTTGTAGTCTATATTTGATA 1871
 |||||
 Db 851 TAT 792
 |||||
 Qy 1872 ATTTGATATATGATCAAGTCTCTAACTCAATGCTCAATGATATATATATATATATATAT 1931
 |||||
 Db 791 AT 732
 |||||
 Qy 1932 T 1932
 |||||
 Db 731 T 731

RESULT 10

US-09-949-016-15940/c
 ; Sequence 15940, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15940
 ; LENGTH: 191569
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(191569)
 ; OTHER INFORMATION: n = A, T, C or G
 US-09-949-016-15940

Query Match 3.2%; Score 83.8; DB 3; Length 191569;
 Best Local Similarity 49.6%; Pred. No. 5.8e-08;
 Matches 298; Conservative 0; Mismatches 297; Indels 6; Gaps 3;
 Qy 1332 TTTAAGTTACCAACATTTTGTGAATTAATTTGGCATATATGATTAATACCAACATATTT 1391
 |||||
 Db 1325 TTTATTTTATATATATATTTTATTTTATATATATATATATATATATATATATATATAT 1266
 |||||
 Qy 1392 ATATGCGAATTTTAAATCTATACGATGATACGGTTTATGGAGTAGACCGCAAAAT 1451
 |||||
 Db 1265 TATATAAATATATATTTTATATATATATATAAATATATATATATATATATATATATA 1206
 |||||
 Qy 1452 ATTATGTATACGGAATGCAATGAGATAGATAATACATTTTGTGGGCTCTTTCGACTT 1511
 |||||
 Db 1205 TATTTA 1146
 |||||
 Qy 1512 ATATGCGTCACCATTTTGAACCAATAAATTTTAAATTTTCTATGATATATATATATAT 1571
 |||||
 Db 1145 ATATATATATATATATATATATATATA-TTATATATTTTATATATATATATATATAT 1087
 |||||
 Qy 1572 TATGATGTCATGCAATGACAGCTAAACACAGGTTTGACATATATATCTATGTTATGT 1631
 |||||
 Db 1086 TATATATATATATATTTTATACATATATATATATATATATATATATATATATATAT 1027
 |||||
 Qy 1632 ATTGCACATTCACCTTGACTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1691
 |||||
 Db 1026 TTTAT 968
 |||||

Qy 1692 TAAATATAATCACTTAATTTATATATTTTGCATTTGTACCTCATATCATAGTACTCTAAATTT 1751
 |||||
 Db 967 ATATATTTTAT 912
 |||||
 Qy 1752 CTCTCAACGCTGCTATCTTTTGTGCTAACTGCTAACTTTACATAGTTTGTGCAATCTTCTT 1811
 |||||
 Db 911 TTTAT 852
 |||||
 Qy 1812 TCAAAACCATATCTTCGATAAATGATATTTTTCATAGATATTTGTAGTCTATATTTGATA 1871
 |||||
 Db 851 TAT 792
 |||||
 Qy 1872 ATTTGATATATGATCAAGTCTCTAACTCAATGCTCATATATATATATATATATATAT 1931
 |||||
 Db 791 AT 732
 |||||
 Qy 1932 T 1932
 |||||
 Db 731 T 731

RESULT 11

US-09-889-926-1
 ; Sequence 1, Application US/09889926
 ; Patent No. 6844486
 ; GENERAL INFORMATION:
 ; APPLICANT: Xie, Qi
 ; APPLICANT: Chua, Nam-Hai
 ; APPLICANT: Institute of Molecular Agrobiolgy, The National U
 ; TITLE OF INVENTION: NAC1 - A PLANT GENE ENCODING A TRANSCRIPTION FACTOR
 ; INVOLVED IN COTYLEDON AND LATERAL ROOT DEVELOPMENT
 ; FILE REFERENCE: 2248-115
 ; CURRENT APPLICATION NUMBER: US/09/889,926
 ; CURRENT FILING DATE: 2001-08-18
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1287
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (89)...(1060)
 US-09-889-926-1

Query Match 3.1%; Score 81.4; DB 3; Length 1287;
 Best Local Similarity 61.3%; Pred. No. 5.2e-08;
 Matches 149; Conservative 0; Mismatches 91; Indels 3; Gaps 1;
 Qy 620 CAGCTATGGCGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAT 679
 |||||
 Db 285 CAAAAATGGCATCGGTGGGAGGAAGGATTTGTTTCTACAGCCAAAGAGACCGAAAT 344
 |||||
 Qy 680 ATAGAATGAGATAGACCGAAGAGTACAGCTTCCAGGATATTTGAAAGCCACCGAG 739
 |||||
 Db 345 ACCGACGGGCTGAGAACTTAACCGAGACCGGCCCGGATTTGGAAGCCACCGCA 404
 |||||
 Qy 740 CTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAAACCTAGTTT 799
 |||||
 Db 405 AAGACAGAACCAITCTTAAGAAAG--GGTAAAGCTAGTTGGGATGAGAGACATTTGTTT 461
 |||||
 Qy 800 TCTACTCTGGTAAAGCCCTTAAAGGCACTCGTACTAGTTGGATCATCAAGAGTATCGTC 859
 |||||
 Db 462 TCTATCAAGGTCGAGCTCTCGAGCGCGTAAACCCGATTTGGGTCTATGACGAAATTCGTC 521
 |||||
 Qy 860 TTC 862
 |||||
 Db 522 TCC 524

RESULT 12

US-09-949-016-17145/c

Db 2341 TATCAGATGGAGGTAAACATCAATCAACAAATCTTTTCAATTTGCTCAACAGTTTTCATCATA 2400
Qy 2401 CTCACACAAATGCTTAACGCAACGCAATTAACATTTGGTGGCTGGCGGCACTACAGCGA 2460
Db 2401 CTCACACAAATGCTTAACGCAACGCAATTAACATTTGGTGGCTGGCGGCACTACAGCGA 2460
Qy 2461 CAACGCTAAATGCTTCAAACTCAAGCGGGGTAGCTATGAACATGATTTCTTCAGGAACGA 2520
Db 2461 CAACGCTAAATGCTTCAAACTCAAGCGGGGTAGCTATGAACATGATTTCTTCAGGAACGA 2520
Qy 2521 TTCCAAACAATGCTTTGGGATATGTGGATCCAAATAGTACAGATGGAACAGAGATC 2580
Db 2521 TTCCAAACAATGCTTTGGGATATGTGGATCCAAATAGTACAGATGGAACAGAGATC 2580
Qy 2581 ACTATACTAATATCTCTTTTAAAGTAA 2606
Db 2581 ACTATACTAATATCTCTTTTAAAGTAA 2606

RESULT 2

US-10-225-066A-115

; Sequence 115, Application US/10225066A

; Publication No. US20030226173A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: RATCLIFFE, Oliver

; APPLICANT: RIECHMANN, Jose Luis

; APPLICANT: ADAM, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omalra

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROUN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: M810316-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR FILING DATE: 2002-08-09

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-08-09

; PRIOR FILING DATE: 2001-08-09

; PRIOR FILING DATE: 2001-12-05

; PRIOR FILING DATE: 2001-12-11

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 1122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 115

; LENGTH: 1444

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; US-10-225-066A-115

; Query Match 26.4%; Score 687.2; DB 6; Length 1444;

; Best Local Similarity 96.8%; Pred. No. 1e-124;

; Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1883 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATAGGCTGAATATCATTTGTCGCG 1942

Db 669 GTATCGTCTTCCGCAACCATGAACCGAGAGTACCAAAAGGCTGAATATCATTTGTCGCG 728

Qy 1943 AGTGTAACAAAGGCGAGAGTAGAAGATCATCCATCGGTACCGATTTCTCTCCCAAG 2002

Db 729 AGTGTAACAAAGGCGAGAGTAGAAGATCATCCATCGGTACCGATTTCTCTCCCAAG 788

Qy 2003 ACATCATAACCATCACTCATCGATCATCCGTTAGCCTTAGACAAACACACCA 2062

Db 789 ACATCATAACCATCACTCATCGATCATCCGTTTAGCCTTAGACAAACACACCA 848
Qy 2063 TTCTATCTCTCTTAATCAATTCGACACAACTTAAACAAACAAACATCAACATCT 2122
Db 849 TTCTATCTCTCTTAATCAATTCGACACAACTTAAACAAACAAACATCAACATCT 908
Qy 2123 CGAAGAGCTCTCCACCGAATATTTCCGGCGAGCGGAGCAGCAACAAACAGCAACAAACAG 2182
Db 909 CGAAGAGCTCTCCACCGAATATTTCCGGCGAGCGGAGCAGCAACAAACAGCAACAAACAG 968
Qy 2183 TAACTCTGAGCTTCAATTTGCTCTAGCCAATCAAAAATATATATGTCCTTACGA 2242
Db 969 TAACTCTGAGCTTCAATTTGCTCTAGCCAATCAAAAATATATATGTCCTTACGA 1028
Qy 2243 CACAAGCAACACACATTTGATGCTCTACGAGAAATCATCAAGACGATGATGAACCTGC 2302
Db 1029 CACAAGCAACACACATTTGATGCTCTACGAGAAATCATCAAGACGATGATGAACCTGC 1088
Qy 2303 CATTGTTGAGCTCTTCAAAAGACTAGTTAACTACCAATATCAGATGGAGGTAACTCAA 2362
Db 1089 CATTGTTGAGCTCTTCAAAAGACTAGTTAACTACCAATATCAGATGGAGGTAACTCAA 1148
Qy 2363 TCACCAATCTTTCAAATTTGCTCAACAGTTTTCATATCTCAACAAACAAATGCTAACGC 2422
Db 1149 TCACCAATCTTTCAAATTTGCTCAACAGTTTTCATATCTCAACAAACAAATGCTAACGC 1208
Qy 2423 AAACGCAATTACAAATTTGCTGGCGGACTACAGGACAACTGCTCAAACTCA 2482
Db 1209 AAACGCAATTACAAATTTGCTGGCGGACTACAGGACAACTGCTCAAACTCA 1268
Qy 2483 AGCGGGCTTAGCTATGAACATGATTTCTGCGAGAAACGATTTCCAAACAAATGCTTTGTGGGA 2542
Db 1269 AGCGGGCTTAGCTATGAACATGATTTCTGCGAGAAACGATTTCCAAACAAATGCTTTGTGGGA 1328
Qy 2543 TATGTGGAATCCAATAGTACCAAGTGAACAGAGATCACTATATCTTCTTTAA 2602
Db 1329 TATGTGGAATCCAATAGTACCAAGTGAACAGAGATCACTATATCTTCTTTAA 1388
Qy 2603 GTAA 2606
Db 1389 GTAA 1392

RESULT 3

US-10-225-067-87

; Sequence 87, Application US/10225067

; Publication No. US20040019925A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: Heard, Jacqueline E.

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Keddle, James

; APPLICANT: Pilgrim, Marsha L.

; APPLICANT: Dubell, Arnold T.

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pineda, Omalra

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Broun, Pierre E.

; TITLE OF INVENTION: BIOCHEMISTRY-RELATED POLYNUCLEOTIDES AND

; FILE REFERENCE: 51442002042

; CURRENT APPLICATION NUMBER: US/10/225,067

; CURRENT FILING DATE: 2002-08-09

; PRIOR FILING DATE: 2002-08-09

; PRIOR FILING DATE: 2001-08-09

; PRIOR FILING DATE: 2001-08-09

; PRIOR FILING DATE: 2001-11-19

; PRIOR FILING DATE: 2001-12-11

; PRIOR FILING DATE: 2002-06-14

```

; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)...(1392)
US-10-225-067-87

Query Match
Best Local Similarity 26.4%; Score 687.2; DB 7; Length 1444;
Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1883 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATTAGGCTGAATATCATTTGCGG 1942
Db 669 GTATCGTCTTCGGACCATGAAACCGAGAAGTACCAAAAGGCTGAAATATCATTTGCGG 728

QY 1943 AGTGTAACAAAGGCGAGGAGTAGATCATCCATCGGTACCAGCTTCTCTCCACAAAG 2002
Db 729 AGTGTAACAAAGGCGAGGAGTAGATCATCCATCGGTACCAGCTTCTCTCCACAAAG 788

QY 2003 ACATCATCAACCAATCACTCATCGACATCATCCCGTTTAGCTTTAAGACAAACAAACCA 2062
Db 789 ACATCATCAACCAATCACTCATCGACATCATCCCGTTTAGCTTTAAGACAAACAAACCA 848

QY 2063 TTATCTCTCTCTAATCAATTCGACAAACCTTAACAAACAAACAAACATCAACAATCT 2122
Db 849 TTATCTCTCTCTAATCAATTCGACAAACCTTAACAAACAAACAAACATCAACAATCT 908

QY 2123 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGACGACAAACAAACAAACGACAAACAG 2182
Db 909 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGACGACAAACAAACAAACGACAAACAG 968

QY 2183 TAACCTCTGACGTTACATTTGCTCTAGCAATCAAAACATATATCTGCAATGCTTACGA 2242
Db 969 TAACCTCTGACGTTACATTTGCTCTAGCAATCAAAACATATATCTGCAATGCTTACGA 1028

QY 2243 CACAAGCAACCAACATGATGTCTCTACGAGAAATCATCAAGACGATGATGAACATGC 2302
Db 1029 CACAAGCAACCAACATGATGTCTCTACGAGAAATCATCAAGACGATGATGAACATGC 1088

QY 2303 CATTTGACGATCTTCAAGACTAGTTAACTACCAATATCAGATGAGGTAAACATCA 2362
Db 1089 CATTTGACGATCTTCAAGACTAGTTAACTACCAATATCAGATGAGGTAAACATCA 1148

QY 2363 TCACCAATACTTTCAAAATTTGCTCAACAGTTTTCATCATCTCAACAAACAAATGCTAAACGC 2422
Db 1149 TCACCAATACTTTCAAAATTTGCTCAACAGTTTTCATCATCTCAACAAACAAATGCTAAACGC 1208

QY 2423 AAACGCATTACAAATTTGGTGGCTGCGGAGTACAGCAACGCTTAATGCTTCAAACTCA 2482
Db 1209 AAACGCATTACAAATTTGGTGGCTGCGGAGTACAGCAACGCTTAATGCTTCAAACTCA 1268

QY 2483 AGCGGGTTAGTATGAACATGATTTCTCGAGGACGATTCACAAATGCTTTGTGGGA 2542
Db 1269 AGCGGGTTAGTATGAACATGATTTCTCGAGGACGATTCACAAATGCTTTGTGGGA 1328

QY 2543 TATGTGGAATCCAATAGTACGAGATGGAACAGAGATCACTATCTAATATTTCTTTTAA 2602
Db 1329 TATGTGGAATCCAATAGTACGAGATGGAACAGAGATCACTATCTAATATTTCTTTTAA 1388

QY 2603 GTAA 2606
Db 1389 GTAA 1392

```

```

RESULT 4
US-10-374-780A-2475
; Sequence 2475, Application US/10374780A
; Publication No. US2004001927A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2475
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G962
US-10-374-780A-2475

Query Match
Best Local Similarity 26.4%; Score 687.2; DB 7; Length 1444;
Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1883 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATTAGGCTGAATATCATTTGCGG 1942
Db 669 GTATCGTCTTCGGACCATGAAACCGAGAAGTACCAAAAGGCTGAAATATCATTTGCGG 728

QY 1943 AGTGTAACAAAGGCGAGGAGTAGATCATCCATCGGTACCAGCTTCTCTCCACAAAG 2002
Db 729 AGTGTAACAAAGGCGAGGAGTAGATCATCCATCGGTACCAGCTTCTCTCCACAAAG 788

QY 2003 ACATCATCAACCAATCACTCATCGACATCATCCCGTTTAGCTTTAAGACAAACAAACCA 2062
Db 789 ACATCATCAACCAATCACTCATCGACATCATCCCGTTTAGCTTTAAGACAAACAAACCA 848

QY 2063 TTATCTCTCTCTAATCAATTCGACAAACCTTAACAAACAAACAAACATCAACAATCT 2122
Db 849 TTATCTCTCTCTAATCAATTCGACAAACCTTAACAAACAAACAAACATCAACAATCT 908

QY 2123 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGACGACAAACAAACAAACGACAAACAG 2182
Db 909 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGACGACAAACAAACAAACGACAAACAG 968

QY 2183 TAACCTCTGACGTTACATTTGCTCTAGCAATCAAAACATATATCTGCAATGCTTACGA 2242

```

Db 969 TAACTCTGACGTTACCATTTGCTCTAGCCAAATCAAAACATATATCGTCCAATGCGCTTACGA 1028
 Qy 2243 CACAGCAACACACATTTGATGCTCTACGAGAAATCATCAAGACGATGATGAACCTGC 2302
 Db 1029 CACAGCAACACACATTTGATGCTCTACGAGAAATCATCAAGACGATGATGAACCTGC 1088
 Qy 2303 CATTGTTGACGATCTTCAAGACTAGTTAACTTACCAATATACAGATGAGGTAAACATCAA 2362
 Db 1089 CATTGTTGACGATCTTCAAGACTAGTTAACTTACCAATATACAGATGAGGTAAACATCAA 1148
 Qy 2363 TCACCAATATCTTCAAAATGCTTCAACAGTTTCACTCATATCACTCAAAACAAATGCTTAACGC 2422
 Db 1149 TCACCAATATCTTCAAAATGCTTCAACAGTTTCACTCATATCACTCAAAACAAATGCTTAACGC 1208
 Qy 2423 AAACGCTTACCAATTTGGTGGCTGGCGGACTACAGCGAACGCTTAATGCTTCAAACTCA 2482
 Db 1209 AAACGCTTACCAATTTGGTGGCTGGCGGACTACAGCGAACGCTTAATGCTTCAAACTCA 1268
 Qy 2483 AGCGGCTTACGATGAAACATGATTTCTGCGAGAACGATTTCAAAACAAATGCTTGTGGGA 2542
 Db 1269 AGCGGCTTACGATGAAACATGATTTCTGCGAGAACGATTTCAAAACAAATGCTTGTGGGA 1328
 Qy 2543 TATGTGGAATCCAATAGTACGAGATGGAACAGAGATCACTATATATTCCTTTTAA 2602
 Db 1329 TATGTGGAATCCAATAGTACGAGATGGAACAGAGATCACTATATATTCCTTTTAA 1388
 Qy 2603 GTAA 2606
 Db 1389 GTAA 1392

RESULT 5

US-10-225-066A-115
 ; Sequence 115, Application US/10225066A
 ; Publication No. US20050160493A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E
 ; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROWN, Pierre E
 ; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: ME10036-2 US
 ; CURRENT APPLICATION NUMBER: US/10/225,066A
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/310,847
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/336,049
 ; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 115
 ; LENGTH: 1444
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-225-066A-115
 Query Match 26.4%; Score 687.2; DB 9; Length 1444;
 Best Local Similarity 96.8%; Pred. No. 1e-124;

Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 Qy 1983 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATTATAGCTGAATATCATTTGTGCG 1942
 Db 669 GTATCGTCTTCCGACCATGAAACCGAGAGTAGTCCAAAGGCTGAATATCATTTGTGCG 728
 Qy 1943 AGTGTACAAAAGGCGAGGAGTAGAAGATCATCCATCGGTACACGTTCTCTCTCCACAAG 2002
 Db 729 AGTGTACAAAAGGCGAGGAGTAGAAGATCATCCATCGGTACACGTTCTCTCTCCACAAG 788
 Qy 2003 ACATCATAAACATTAATCATATCGACATCATCCGTTTACGCTTAAAGACAAACACACA 2062
 Db 789 ACATCATAAACATTAATCATATCGACATCATCCGTTTACGCTTAAAGACAAACACACA 848
 Qy 2063 TTCACTCTCTCTAATCATTTCCGACAAACCTTTAAACAAACAAACATCAACAATCT 2122
 Db 849 TTCACTCTCTCTAATCATTTCCGACAAACCTTTAAACAAACAAACATCAACAATCT 908
 Qy 2123 CGAGAAGCTCTCCACCGAATATTCCGGCGAGCGAGCAGCAACAAACAGCAACACACAG 2182
 Db 909 CGAGAAGCTCTCCACCGAATATTCCGGCGAGCGAGCAGCAACAAACAGCAACACAG 968
 Qy 2183 TAATCTGAGCTTACCAATTTGCTCTAGCAATCAAAAATATATATCGTCAATGCTTACGA 2242
 Db 969 TAATCTGAGCTTACCAATTTGCTCTAGCAATCAAAAATATATATCGTCAATGCTTACGA 1028
 Qy 2243 CACAGCAACACACATTTGATGCTCTACGAGAAATCATCAAGACGATGATGAACCTGC 2302
 Db 1029 CACAGCAACACACATTTGATGCTCTACGAGAAATCATCAAGACGATGATGAACCTGC 1088
 Qy 2303 CATTGTTGACGATCTTCAAAAGACTAGTTAACTACCAAAATATCAATGAGGTAACTCAA 2362
 Db 1089 CATTGTTGACGATCTTCAAAAGACTAGTTAACTACCAAAATATCAATGAGGTAACTCAA 1148
 Qy 2363 TCACCAATATCTTCAAAATGCTCTCAACAGTTTCAATCATATCAACAAACAAATGCTTAACGC 2422
 Db 1149 TCACCAATATCTTCAAAATGCTCTCAACAGTTTCAATCATATCAACAAACAAATGCTTAACGC 1208
 Qy 2423 AAACGCTTACCAATTTGGTGGCTGGCGGACTACAGCGAACGCTTAATGCTTCAAACTCA 2482
 Db 1209 AAACGCTTACCAATTTGGTGGCTGGCGGACTACAGCGAACGCTTAATGCTTCAAACTCA 1268
 Qy 2483 AGCGGCTTACGATGAAACATGATTTCTGCGAGAACGATTTCAAAACAAATGCTTGTGGGA 2542
 Db 1269 AGCGGCTTACGATGAAACATGATTTCTGCGAGAACGATTTCAAAACAAATGCTTGTGGGA 1328
 Qy 2543 TATGTGGAATCCAATAGTACGAGATGGAACAGAGATCACTATATATTCCTTTTAA 2602
 Db 1329 TATGTGGAATCCAATAGTACGAGATGGAACAGAGATCACTATATATTCCTTTTAA 1388
 Qy 2603 GTAA 2606
 Db 1389 GTAA 1392

RESULT 6

US-10-780-703-1
 ; Sequence 1, Application US/10780703
 ; Publication No. US20050034194A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, Jong Seob
 ; APPLICANT: Kim, Yun Hee
 ; APPLICANT: Choi, Eun Kyung
 ; APPLICANT: Yoo, So Yeon
 ; APPLICANT: Ahn, Ji Hoon
 ; APPLICANT: Choi, Yang Do
 ; TITLE OF INVENTION: Gene Controlling Flowering Time of Plants and Method for
 ; FILE REFERENCE: 012679-105
 ; CURRENT APPLICATION NUMBER: US/10/780,703
 ; CURRENT FILING DATE: 2004-02-19
 ; PRIOR APPLICATION NUMBER: KR 10-2003-10772
 ; PRIOR FILING DATE: 2003-02-20

; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1140
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)....(1137)
 ; OTHER INFORMATION: cDNA of LOV1 gene
 US-10-780-703-1

Query Match		17.9%;	Score 467.2;	DB 8;	Length 1140;
Best Local Similarity		82.3%;	Pred. No. 1.6e-81;		
Matches 596;		Conservative 0;	Mismatches 23;	Indels 105;	Gaps 1;

QY	1883	GTATCAAGTCTCTAATCAATGTCTCATGTATAATATATAGTGTGAATATCATTTGTGCGG	1942
DB	522	GTATCGTCTTCGCGACCATGAAACCGAAGATACCAAAAGGCTGAAATATCATTTGTGCGG	581
QY	1943	AGTGTACAAAAGGCCAGGAGTAGAAGATCATTCATCGGTACCAAGTCTCTCTCCACAA	2002
DB	582	AGTGTACAAAAGGCCAGGAGTAGAAGATCATTCATCGGTACCAAGTCTCTCTCCACAA	641
QY	2003	ACATCATACCAATACTCATCGACATCATCCGTTTGTAGCTTAAAGCAACAACAACACCA	2062
DB	642	ACATCATACCAATACTCATCGACATCATCCGTTTGTAGCTTAAAGCAACAACAACACCA	701
QY	2063	TTTCATCTCTCTCTAATCAATTCGACAAACCTTTAAACAACAACAACAACAATCT	2122
DB	702	TTTCATCTCTCTCTAATCAATTCGACAAACCTTTAAACAACAACAACAACAATCT	761
QY	2123	CGAAGAGCTCTCCACGAATATTCGCGGACGCGCAGCACAACAACAACAACAACAAG	2182
DB	762	CGAAGAGCTCTCCACGAATATTCGCGGACGCGCAGCACAACAACAACAACAACAAG	821
QY	2183	TAACTCTGACGTTACCATGCTCTAGCCTCAATCAAAACATATATCGTCCAAATGCTTACGA	2242
DB	822	TAACTCTGACGTTACCATGCTCTAGCCTCAATCAAAACATATATCGTCCAAATGCTTACGA	881
QY	2243	CACAAGCAACAACAATGATGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC	2302
DB	882	CACAAGCAACAACAATGATGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC	941
QY	2303	CATTGTGACGATCTTCAAGACTAGTTAACTACCAAAATATCATGATGGAGGTAAACATCA	2362
DB	942	CATTGTGACGATCTTCAAGACTAGTTAACTACCAAAATATCATGATGGAGGTAAACATCA	989
QY	2363	TCACCAATACTTTCAAATTGCTCAACAGTTTCATCATCTACACACAACAATGCTTAACGC	2422
DB	990	TCACCAATACTTTCAAATTGCTCAACAGTTTCATCATCTACACACAACAATGCTTAACGC	989
QY	2423	AAACGATTACAATTGGTGGCTGCGGAGTACAGCGCAACGCTAATGCCTCAAACTCA	2482
DB	990	AAACGATTACAATTGGTGGCTGCGGAGTACAGCGCAACGCTAATGCCTCAAACTCA	1016
QY	2483	AGCGCGTTAGCTATGAACATGATTCCTCGAGGAACGATTCACAAATGCTTTGTGGGA	2542
DB	1017	AGCGCGTTAGCTATGAACATGATTCCTCGAGGAACGATTCACAAATGCTTTGTGGGA	1076
QY	2543	TATGTGGAATCCATAGTACCAATGATGGAACAGAGATCACTACTAATATTCCTTTAA	2602
DB	1077	TATGTGGAATCCATAGTACCAATGATGGAACAGAGATCACTACTAATATTCCTTTAA	1136
QY	2603	GTAA 2606	
DB	1137	GTAA 1140	

RESULT 7
 US-10-424-599-121054
 ; Sequence 121054, Application US/10424599
 ; Publication No. US20040031072A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 121054
 ; LENGTH: 785
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_80317C.1
 US-10-424-599-121054

Query Match		7.1%;	Score 184.4;	DB 7;	Length 785;
Best Local Similarity		78.4%;	Pred. No. 5.2e-26;		
Matches 221;		Conservative 0;	Mismatches 61;	Indels 0;	Gaps 0;

QY	618	TGCAGCTATGCGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCGCAAGAGATCGGAA	677
DB	358	TCCGCGCTTGGCAGCTATTGGTGAGAGAGTGGTACTTCTATGTGCGCGAGATAGGAA	417
QY	678	ATATAGAATGGAGATAGACCGAACCGAGTAACGATTCAGGATATTGGAAGCCACCGG	737
DB	418	GTATGCAACGGTGTATCGTCCATCGTGTACTACCTCTGGGTATTGGAAGCGAACAGG	477
QY	738	AGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTACT	797
DB	478	AGCTGATAGGATGATCAGGACCGAGAAATTTTCGCTCTATCGCCTCAAGAAACCCCTACT	537
QY	798	TTTCTACTCTGGTAAAGCCCTTAAAGGCACCTCGTACTAGTTGGATCGATGAACGAGTATCG	857
DB	538	TTTCTATTCTGGAAAGCTCTTAAAGGCATACGAACCGATTTGGATTATGAACGAGTATCG	597
QY	858	TCTTCGCGACCATGAAACCGGAGAGTACCAAAAGGTATAAAT	899
DB	598	CTTTCGCGAATGAAACTGAAACGATATCAAAAGGCTGAGAT	639

RESULT 8
 US-10-437-963-88434
 ; Sequence 88434, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 88434
 ; LENGTH: 1433
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_87286C.1
 US-10-437-963-88434

Query Match		6.3%;	Score 164;	DB 7;	Length 1433;
Best Local Similarity		70.8%;	Pred. No. 6.9e-22;		

Matches	218;	Conservative	0;	Mismatches	90;	Indels	0;	Gaps	0
Qy	605	AGATATATGTATATGCAGCTATGGCGGGATAGGAGAGAAAGAGTGGTACTTCTATGTGC	664						
Db	287	AGTCGATGGGGTGCAGCAATGGCGCGATAGGGAGAGGATGGTCTTCTACGTGC	346						
Qy	665	CAAGAGATCGGAATATATAGAAATGGAGATAGACCGAAACCGAGTAAACGACTTCAGGATATT	724						
Db	347	CTCGGGACAGGAAGTACAGGAACCGGGACCGGCCGAAACCGGGTACCGCGTCTCGGGTACT	406						
Qy	725	GGAAAGCCACCGGAGCTATAGGATGTCAGATCGGAGACTTCTCGGCGCTATCGGATTA	784						
Db	407	GGAAAGGCGACGGGGGCCACCCGGATGATCCGAGCCGGAACAACCGCCCGCTCGGGCTCA	466						
Qy	785	AGAAAAACCTAGTTTTTCTACTCTGGTAAAGCCCCCTAAAAGGCACTCGTACTAGTTGGATCA	844						
Db	467	AGAAGACGCTTGTCTTCTACTTCCGCAAGGCCCCCAAGGGCGTCCGCGAGGAGCTGGATCA	526						
Qy	845	TGAACGAGTATCGTCTTCCGCAACCATGAAACCGAGAAAGTACCAAAAGGTATATAATTTCTAC	904						
Db	527	TGAACGAGTATCGCTTCTCCCGCCCGCGACACCGACCGCTACCAACAAGACTGAAATCTCAC	586						
Qy	905	TATAACTC	912						
Db	587	TATGCCG	594						

```

RESULT 9
US-10-437-963-81769
; Sequence 81769, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81769
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(743)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81262C.1
US-10-437-963-81769

```

	Query Match	6.0%	Score 155.6;	DB 7;	Length 743;
	Best Local Similarity	71.0%;	Pred. No. 2.3e-20;		
	Matches 206;	Conservative 0;	Mismatches 84;	Indels 0;	Gaps 0;
Qy	618	TCCGAGCTATCGCGCGATAGAGAGAGAAGAGTGTCTTCTATCTGCCAAGAGATCGGAA	677		
Db	336	TCCGCGCTCTGGCCCTCAATTGGGGACAAGAGTGTCTTCTATCTGCCAAGGAGCCGCAA	395		
Qy	678	ATATAGAAATGGAGATAGACCGAACCAGAGTAAACGACTTCAGGATATTGGAAAGCCACCGG	737		
Db	396	GTACCGAAACGGCGATCGGCCCAACCGAGTAGCGCCATCGGGGTACTGGAAGGCCACAGG	455		
Qy	738	AGCTGATAGGATGATCAGATCGGAGACTTTCGGCCCTATCGGATTTAAAGAAAAACCCCTAGT	797		
Db	456	GGCCGATAGATGCTGAAGGTGGAAGGGGTGACCGTCTTATCGCCTCAAGAAGAAGCGCTCGT	515		

Qy	798	TTTTCTACTCTGGTAAAGCCCTTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCG	857
Db	516	CTTCTACGTGCGCAAGGCGCCCAAGGGGCTTCGACGACGTGATCATGAACGAGTATCG	575
Qy	858	TC TTCCGCACCATGAAGAACCGAGAGTACCAAAAGGTATAAATTCCTACTAT	907
Db	576	TCTCCCTCATGGCGACGTGACCGCTACCAAAAGGTAGAAAATTCCTGCTGT	625

```

RESULT 10
US-10-425-115-101230
; Sequence 101230, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 101230
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(390)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_23835C.1
US-10-425-115-101230

```

Query Match	5.6%	Score 145.2;	DB 8;	Length 390;
Best Local Similarity	69.6%;	Pred. No. 1.9e-18;		
Matches 195;	Conservative 0;	Mismatches 85;	Indels 0;	Gaps 0;
Qy	620	CAGCTATGGCGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAT	679	
Db	55	CATCAATGGCGCGATTGGGGGAAGAGGGGTAGTTCTACGTCCCGAGGACCGCAAGT	114	
Qy	680	ATAGAAATGGAGATAGACCGAAACCGAGTACAGCTTCAGGATATTTGAAAGCCACCGGAG	739	
Db	115	ACGCGAACGGCGACAGGCCCAACGGGTACGGGTCTGGGCTACTGGAAGCCACGGCGC	174	
Qy	740	CTGATAGGATGATCAGATCGGAGACTTCTCGGGCTATCGGATTAAGAAAAACCTAGTTT	799	
Db	175	CCGACCGCACCATCAAGGCGGACCAACAGCCGCCCATCGGGCTCAAGAACGCTCGTCT	234	
Qy	800	TCTACTCTGTAAAGCCCTTAAGGCACCTCGTACTAGTTGGATCATGAACGAGTATCGTC	859	
Db	235	TCTACTCCGCAAGGCGCCCAAGGGGTGCGCAGCAGCTGGATCATGAACGAGTACCGCC	294	
Qy	860	TTCCGACCATGAAACCGAGAAGTACCAAAAGGTATAAAT	899	
Db	295	TCCCGGCCGACGACACCGACCGTGTACCAAGAACCGAAAT	334	

RESULT 11
US-10-425-115-173486
; Sequence 173486, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 173486
LENGTH: 792
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_897C.1
US-10-425-115-173486

Query Match
Best Local Similarity 5.4%; Score 140; DB 8; Length 792;
Matches 191; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 618 TCAGCTATGGCGGCGATAGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAA 677
DB 417 TCCAGCAATGGCGGTGATGGCGGGAAGAGTGGTCTTCTACGTGCCGAGGACCGTAA 476
QY 678 ATATAGAAATGAGATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAGCCACCGG 737
DB 477 GTACCGGAACGAGACCGCCCAACCGGGTGACGGCTCGGGGTACTGGAAGCGACCGG 536
QY 738 AGCTGATAGATGATCAGATCGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGT 797
DB 537 CCGTACCGGATGATCCGAGGCGAGAACACCGCCCATCGGCTGAGAGACGCTCGT 596
QY 798 TTCTACTCTGTAAGCCCTTAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857
DB 597 GTTCTACTCCGACGAGCGCCCAAGGGCTCGCAGCAGCTGATCATGAACGAGTACCG 656
QY 858 TCTTCGCGACCATGAACCGGAGAGTACCAAAAGT 893
DB 657 CTTCCGCGCACCCACCGACCGCGCCGACCCATTGAT 692

RESULT 12
US-10-425-115-10100
Sequence 10100, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 10100
LENGTH: 1291
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_109212C.1
US-10-425-115-10100

Query Match
Best Local Similarity 5.0%; Score 131.2; DB 8; Length 1291;
Matches 196; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 620 CAGCTATGGCGGCGATAGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAT 679
DB 286 CTGCTCTGCTTCGATCGGGACAGAGTGGTACTTCTACGTGCCGAGAGACCGCACT 345
QY 680 ATAGAAATGAGATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAGCCACCGGAG 739
DB 346 ACCGCAACGGGACCGCCCAACCGGGTCAACCGGCTACGCGCTACTGGAAGCGACGGGG 405
QY 740 CTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTT 799

DB 406 CCGACCGCACGGTCTACGTGAGGTCAAGCGCGCGCTCAAGAGACGCTGTGT 465
QY 800 TCTACTCTGGTAAAGCCCTTAAAGGCACTCTGACTAGTTGGATCATGAACGAGTATCGTC 859
DB 466 TCTACGTGGGAGAGCGGCCCAAGGACTCAGGAGCAGCTGATCATGAACGAGTACCGCC 525
QY 860 TTCCGCACCATGAACCGAGAGTACCAAAAGGTATAAATCTACTATAACTCTATATAT 919
DB 526 TTCCTTCAGCGAGCGCGACCGCTACCAAAAGGAAATTTCTTTTGGAAAGTGTACAAC 585
QY 920 ATCC 923
DB 586 GACC 589

RESULT 13
US-10-424-599-52441
Sequence 52441, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 52441
LENGTH: 1640
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_18368C.1
US-10-424-599-52441

Query Match
Best Local Similarity 4.4%; Score 115; DB 7; Length 1640;
Matches 176; Conservative 0; Mismatches 80; Indels 3; Gaps 1;
QY 609 ATATGTATATCAGCTATGGCGGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAG 668
DB 286 ATGGTCTTCCAGTAAAGGACGCTTTGGAGAGAGAGTGGTACTTCTTCTCAGTCCAAAG 345
QY 669 AGATCGGAAATATAGAAATGGAGATAGACCGAACCGAGTAAACGACTTCAGGATATTGGAA 728
DB 346 AGACAGGAAGTACCCGATGTTCCAGCACCAACACAGATTGCGGGTCTCTGGGTATTGGAA 405
QY 729 AGCACCGGAGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAA 788
DB 406 AGCCACTGGAATGACAAATCATCACCACTGA---AGGTAGAAAAGTTGGCATAAAAA 462
QY 789 AACCTAGTTTCTACTCTGTTAAAGCCCTTAAAGGCACTCGTACTAGTTGGATCATGAA 848
DB 463 AGCACTTGTCTTCTAGCTTGGCAAGACCCCAAGGCTCCAAACCAATTGGATCATGCA 522
QY 849 CGAGTATCGTCTTCGCGAC 867
DB 523 CGAGTATCGCTTCTCGAC 541

RESULT 14
US-10-425-114-29320
Sequence 29320, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E


```
QY 657 CTATGTGCCAGAGATCGGAAATATAGAAATGAGATAGACCGAACCGAGTAACGACTTC 716
DB 371 CTTTAGTCCAGAGATCGGAAGTATCCAAACGAGCAAGACCAACAGAGCGCGGACTTC 430
QY 717 AGGATATTGGAAGCCACCGAGCTGATA---GGATGATCAGATCGGAGACTTCTCGGCC 773
DB 431 AGTTTATGGAAGCCGCGTACAGATAAACCGGTACTTGCTTCGACGGTAACCAAAA 490
QY 774 TATCGGATTAAAGAAAACCTAGTTTCTACTCTGTGTAAGCCCTTAAAGGCACCTCGTAC 833
DB 491 GGTGGGCGTGAAGAGCACTAGTCTTCTACAGTGTGTAACCAACCAAGAGGCGTTAAAG 550
QY 834 TAGTTGGATCATGAACGAGTATCGTCTTCCGACCATGAACCGAGAA 881
DB 551 TGATTGGATCATGCATGAGTATCGTCTCATCGAAGAAACAAACCAACAA 598

RESULT 2
US-11-096-568A-6578
; Sequence 6578, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6578
; LENGTH: 1298
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1298)
; OTHER INFORMATION: Ceres Seq. ID no. 14315846
US-11-096-568A-6578

Query Match
Best Local Similarity 3.5%; Score 92.4; DB 11; Length 1298;
Matches 147; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 641 AGAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAATATAGAAATGAGATAGACCGA 700
DB 368 AGAATGAGTGGTATTTCTTTCAGCCATAAGGACAAAGAAATACCCCAACAGGAACCAACAA 427
QY 701 ACCGAGTAACGACTTCAGGATATTGGAAAGCCCGAGCTGATAGATGATCAGATCGG 760
DB 428 ATCGAGCAACTACAGCTGGTGTGGAAAGCAACTGGAAGAGACAAAGTCCATATACCA 487
QY 761 AGACTTCTCGCCCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGTGTAAGCCCTTA 820
DB 488 CTAATTCAGAGATGTCATGAGGAAACCCCTAGTTTCTACTCTGTGTCGTCCTC 547
QY 821 AAGGCACTCGTACTAGTTCATGAGCAACGAGTATCGTCTTCGCAACCAACCGA 878
DB 548 ATGCCAGAGACTGATTGGATCATGATGATCGCTTGTGATGAAGATGATGCTGA 605

RESULT 3
US-11-096-568A-31996
; Sequence 31996, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
```

```
; SEQ ID NO 31996
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Ceres Seq. ID no. 13591899
US-11-096-568A-31996

Query Match
Best Local Similarity 3.5%; Score 90.8; DB 11; Length 1219;
Matches 162; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

QY 599 AATCATAGATATATATATGATGCGGCGATAGGAGAGAAAGAGTGGTACTTCT 658
DB 197 AATTCGATCCATGGGTTTTTACCAATNAAGCATTTATTTGGAGAAAAGAAATGGTATTTT 256
QY 659 ATGTCCCAAGAGATCGGAAATATAGAAATGAGATAGACCGAACCGAGTAACGACTTCAG 718
DB 257 TTAGTCTTAGGATAGAAATATCCAAACGGGTCAAGACCTAACCGGTTGCGGATCGG 316
QY 719 GATATTGGAAGCCACCGAGCTGATAGATGATCAGATCGGAGACTTCTCGGCCTATCG 778
DB 317 GTTATTGGAAGCTACGGGTACGGATAAAATAATC---TCGACGGAAGGACAAAGAGTTG 373
QY 779 GATTAAAGAAAACCTAGTTTCTACTCTGTGTAAGCCCTTAAAGGCACCTCGTACTAGTT 838
DB 374 GTATTAAAAAAGCTTTGGTGTGTTTACATCGGAAAAGCTCTTAAGGTACTAAAAACCAATT 433
QY 839 GGATCATCAACGAGTATCGTCT 860
DB 434 GGATCATCAAGATGATCGTCT 455

RESULT 4
US-11-121-086-25/c
; Sequence 25, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 173602
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-25

Query Match
Best Local Similarity 3.5%; Score 90.2; DB 14; Length 173602;
Matches 503; Conservative 0; Mismatches 533; Indels 23; Gaps 7;

QY 883 TACCAAAAGGTATAAATCTACTATACTATATATATATATATATATATATATATATATAT 941
DB 137288 TATTATATATATATATATATATATATATATATATATATATATATATATATATATAT 137229
QY 942 TATAACCTTAGCTAGGTGGAGCGCTTTAAATTTAAATTAATTCCTTAGACAGTTTGAA 1001
DB 137228 TAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 137169
QY 1002 TTTTCTTTTGTAGTAGTTTATTTATTTT---TGGATTGATTCGTAAGATCAA 1059
DB 137168 TTATATACATATAATAATATGTTATATAATAATAATAATAATAATAATAATAATAATA 137109
QY 1060 AATACTTGTGATGAGTAAATGTGTCAGCGCGGTTTGGCTTAAATCCAGAAAATGTT 1119
```


LENGTH: 1396

TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1396)
OTHER INFORMATION: Ceres Seq. ID no. 14312047

US-11-096-568A-5850

Query Match 3.3%; Score 86; DB 11; Length 1396;

Best Local Similarity 63.6%; Pred. No. 0.00011;

Matches 150; Conservative 0; Mismatches 80; Indels 6; Gaps 1;

Qy 637 GGAGAGAAGAGTGGTACTTCTATGTCACAGAGATCGGAATATAGAAATCGAGATAGA 696

Db 338 GCGGAGAAGAATGGTACTTCTTCTCTCGGACCGGAAATACCCGACGGTTCACGG 397

Qy 697 CCGAACCGAGTAACGACTTCAGGATATTGGAAGCCACCGAGCTGATAGATATAGA 756

Db 398 CCGAACCGCGCGCGGAGCGGCTATTGGAAGCCACCGCGCTGATAAACCGATCGGA 457

Qy 757 TCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGTTAAAGCC 816

Db 458 AAACCGAAGCGC-----TTGGGATCAAGAAAGCTCTGGTTTTCAGCCGGAAGGCC 511

Qy 817 CCTAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCTTCCGCAACCATGA 872

Db 512 CCCAAGGTGTGAACCAATGGATTATGCAAGATATCGCTTCGCCAATGTTGA 567

RESULT 9

US-11-096-568A-7102

Sequence 7102, Application US/11096568A

Publication No. US20060048240A1

GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.

TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

FILE REFERENCE: 2750-1592PUS2

CURRENT APPLICATION NUMBER: US/11/096.568A

CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 7102

LENGTH: 1395

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)..(1395)

OTHER INFORMATION: Ceres Seq. ID no. 15170403

US-11-096-568A-7102

Query Match 3.2%; Score 83.4; DB 11; Length 1395;

Best Local Similarity 63.0%; Pred. No. 0.00029;

Matches 148; Conservative 0; Mismatches 81; Indels 6; Gaps 1;

Qy 626 TGGCGGCGATAGGAGAGAAGTGGTACTTCTATGTCGCAAGAGATCGGAATATAGAA 685

Db 243 TGGTTCTCTCGGAGAGAAGAGTGGTACTTCTTTTACACCGGACCGGAAATACCGGA 302

Qy 686 ATGAGATAGACCGAACCGAGTAAACGACTTCAGGATATTGGAAGCCACCGGAGCTGATA 745

Db 303 ACGTTTCAGGCCGACCGCGGCGGGAACCGGTTACTGGAAGCAACCGGGGGGATA 362

Qy 746 GGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTTTCTACT 805

Db 363 AGCCCAT-----TGGCCACCCCAACCGGTTGGATATAAAAAAGCTTGGTGTATTACG 416

Qy 806 CTGTTAAGCCCTTAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCT 860

Db 417 CAGGGAAGCTCCGAAGGGGACAAGAGCAATTGGATCATGACGAGTATCGTCT 471

RESULT 10

US-11-096-568A-4700

Sequence 4700, Application US/11096568A

Publication No. US20060048240A1

GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.

TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

FILE REFERENCE: 2750-1592PUS2

CURRENT APPLICATION NUMBER: US/11/096.568A

CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 4700

LENGTH: 1191

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)..(1191)

OTHER INFORMATION: Ceres Seq. ID no. 14303479

US-11-096-568A-4700

Query Match 3.2%; Score 82.6; DB 11; Length 1191;

Best Local Similarity 60.8%; Pred. No. 0.00038;

Matches 155; Conservative 0; Mismatches 94; Indels 6; Gaps 1;

Qy 618 TGCAGCTATGCGCGGATAGGAGAGAAAGAGTGTACTTCTATGTCGCAAGAGATCGAA 677

Db 246 TCAGGATGCGTTTGTACGAGAGAAAGAGTGTATTTTTCACGCGGAGGACCGCA 305

Qy 678 ATATAGAAATGAGATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAGCCACCGG 737

Db 306 GTACCCGACGGTTGCGGCGCGAACCGGTCGCGGGAACCGGTTACTGGAAGGCAACCGG 365

Qy 738 AGCTGATAGATCATGATCGGAGACTTCTCGGCTTATCGGATTAAGAAACCCCTAGT 797

Db 366 AGCGGAT-----AAACCACTTGTGTAACCGGTTGGGATCAAGAAAGCGTTGGT 419

Qy 798 TTCTACTCTGTTAAAGCCCTTAAAGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857

Db 420 TTTTACCTGGAAGAGCGGCCCAAGGAGAGAAACTACTGATCATGACGAGTATCG 479

Qy 858 TCTTCCGACCATGA 872

Db 480 TCTTGCAGACGTGA 494

RESULT 11

US-11-121-086-25

Sequence 25, Application US/11121086

Publication No. US20050266459A1

GENERAL INFORMATION:

APPLICANT: POULSEN, TIM S.

APPLICANT: NIELSEN, KIRSTEN V.

TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

FILE REFERENCE: 09138.6000-00000

CURRENT APPLICATION NUMBER: US/11/121.086

CURRENT FILING DATE: 2005-05-04

PRIOR APPLICATION NUMBER: 60/567,570

PRIOR FILING DATE: 2004-05-04

NUMBER OF SEQ ID NOS: 107

SOFTWARE: PatentIn version 3.3

SEQ ID NO 25

LENGTH: 173602

TYPE: DNA

ORGANISM: Homo sapiens

US-11-121-086-25

Query Match 3.2%; Score 82.2; DB 14; Length 173602;

Best Local Similarity 48.0%; Pred. No. 0.0012;

Matches 393; Conservative 0; Mismatches 413; Indels 13; Gaps 5;

Qy 1126 ATATGCGTGAACCTTTTAAATTTGTAGACATGCGCCCATATGTTATAGTAACATTA 1185

Qy 1828 GATAAATGATATTTTCATAGATATTCCTAGTCTATATTTGATAATTGAT 1878
Db 92 TAGATACAATATATTACATATATAATACAAGTAAATTTATATATAATAAAT 42

Search completed: April 6, 2006, 09:58:15
Job time : 1165 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 10:24:48 ; Search time 580 Seconds
(without alignments)
2614.156 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAVSSTSIIPMSNQVNNN.....MNPDPVGNRDHYTNIPFK 379

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/abs/ABSSWEB.spool/US10780703/runat_06042006_101351_6632/app_query.fasta_1

-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -HOST=abss05h

-USER=US10780703 @CGN_1_1_652 @runat_06042006_101351_6632 -NCPU=6 -ICPU=3

-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMSOUP=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /SIDSS5/ptodata/2/pubpna/US08 NEW PUB.seq.*
2: /SIDSS5/ptodata/2/pubpna/US06 NEW PUB.seq.*
3: /SIDSS5/ptodata/2/pubpna/US07 NEW PUB.seq.*
4: /SIDSS5/ptodata/2/pubpna/PCT NEW PUB.seq.*
5: /SIDSS5/ptodata/2/pubpna/US03 NEW PUB.seq.*
6: /SIDSS5/ptodata/2/pubpna/US09 NEW PUB.seq.*
7: /SIDSS5/ptodata/2/pubpna/US10 NEW PUB.seq.*
8: /SIDSS5/ptodata/2/pubpna/US10 NEW PUB.seq.*
9: /SIDSS5/ptodata/2/pubpna/US10 NEW PUB.seq.*
10: /SIDSS5/ptodata/2/pubpna/US10 NEW PUB.seq.*
11: /SIDSS5/ptodata/2/pubpna/US11 NEW PUB.seq.*
12: /SIDSS5/ptodata/2/pubpna/US11 NEW PUB.seq.*
13: /SIDSS5/ptodata/2/pubpna/US11 NEW PUB.seq.*
14: /SIDSS5/ptodata/2/pubpna/US11 NEW PUB.seq.*
15: /SIDSS5/ptodata/2/pubpna/US60 NEW PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	503	24.7	1263	11	US-11-096-568A-33561 Sequence 33561, A
2	497	24.4	1606	11	US-11-172-740-1242 Sequence 1242, Ap

3	494	24.3	1219	11	US-11-096-568A-31996 Sequence 31996, A
4	488.5	24.0	1191	11	US-11-096-568A-4700 Sequence 4700, Ap
5	482.5	23.7	963	11	US-11-087-099-65 Sequence 65, Appl
6	479.5	23.6	1304	11	US-11-096-568A-7863 Sequence 7863, Ap
7	478	23.5	1396	11	US-11-096-568A-5850 Sequence 5850, Ap
8	474.5	23.3	1423	8	US-10-391-414-4 Sequence 4, Appl
9	473.5	23.3	1395	11	US-11-096-568A-7102 Sequence 7102, Ap
10	471	23.2	1189	11	US-11-096-568A-26116 Sequence 26116, A
11	463	22.8	1409	11	US-11-096-568A-19208 Sequence 19208, A
12	453.5	22.3	1218	11	US-11-087-099-50 Sequence 50, Appl
13	448.5	22.1	1833	11	US-11-096-568A-24275 Sequence 24275, A
14	447.5	22.0	1393	11	US-11-096-568A-22731 Sequence 22731, A
15	444.5	21.9	897	11	US-11-096-568A-30488 Sequence 30488, A
16	441.5	21.7	1316	11	US-11-096-568A-23342 Sequence 23342, A
17	434	21.3	852	11	US-11-096-568A-28895 Sequence 28895, A
18	432	21.2	493	11	US-11-116-881A-790 Sequence 790, App
19	430.5	21.2	1388	11	US-11-096-568A-20504 Sequence 20504, A
20	427.5	21.0	1412	11	US-11-096-568A-3792 Sequence 3792, Ap
21	424.5	20.9	1478	11	US-11-096-568A-19423 Sequence 19423, A
22	423	20.8	1877	11	US-11-096-568A-3133 Sequence 3133, Ap
23	421	20.7	2372	11	US-11-096-568A-24718 Sequence 24718, A
24	419.5	20.6	1192	11	US-11-096-568A-20705 Sequence 20705, A
25	411	20.2	1298	11	US-11-096-568A-6578 Sequence 6578, Ap
26	410.5	20.2	2122	11	US-11-096-568A-31731 Sequence 31731, A
27	403.5	19.8	1528	11	US-11-096-568A-17746 Sequence 17746, A
28	403	19.8	1374	11	US-11-096-568A-13595 Sequence 13595, A
29	400	19.7	918	11	US-11-096-568A-29126 Sequence 29126, A
30	399.5	19.6	1203	11	US-11-096-568A-9631 Sequence 9631, Ap
31	398	19.6	1419	11	US-11-096-568A-31178 Sequence 31178, A
32	397.5	19.5	1251	11	US-11-096-568A-17565 Sequence 17565, A
33	394	19.4	1648	11	US-11-096-568A-25426 Sequence 25426, A
34	391.5	19.2	879	11	US-11-096-568A-30325 Sequence 30325, A
35	388	19.1	596	11	US-11-096-568A-26867 Sequence 26867, A
36	385	18.9	1227	11	US-11-096-568A-34447 Sequence 34447, A
37	380	18.7	1353	11	US-11-096-568A-21255 Sequence 21255, A
38	377.5	18.6	1370	11	US-11-096-568A-17121 Sequence 17121, A
39	372.5	18.3	1285	9	US-10-714-887-261 Sequence 261, App
40	372.5	18.3	1387	11	US-11-096-568A-33784 Sequence 33784, A
41	358.5	17.6	1489	11	US-11-096-568A-22192 Sequence 22192, A
42	350.5	17.2	1421	11	US-11-096-568A-22904 Sequence 22904, A
43	346	17.0	1371	9	US-10-714-887-241 Sequence 241, App
44	345	17.0	1008	11	US-11-096-568A-28163 Sequence 28163, A
45	344.5	16.9	1008	9	US-10-714-887-269 Sequence 269, App

ALIGNMENTS

RESULT 1

US-11-096-568A-33561
; Sequence 33561, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid-
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33561
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1263)
; OTHER INFORMATION: Ceres Seq. ID no. 13603447
US-11-096-568A-33561

Alignment Scores:
Pred. No.: 3.09e-42 Length: 1263
Score: 503.00 Matches: 115
Percent Similarity: 54.8% Conservative: 40

Best Local Similarity: 40.6% Mismatches: 74
Query Match: 24.7% Indels: 54
DB: 11 Gaps: 9

US-10-780-703-2 (1-379) x US-11-096-568A-33561 (1-1263)

```
QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPhePhePheLeuArgArgLys 72
DB 151 CCGGGTTTCGGTTTATCCGACTCAGCAGAGCTGATGGTTGAATATCTCTGTAGAAA 210
QY 73 ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuPheArgPheArg 92
DB 211 GCCGCGGTACGACTTCTCTCCAGCTCATAGCTGAATCGATCTCTACAAGTTTGTAT 270
QY 93 ProThrGluLeuProAlaMetAlaAlaLeuGlyGluLysGluThrPhePheValPro 112
DB 271 CCATCGGTTTACCAAGTAAAGCGTTATTCGGTGAAAGAAATGGTATTTTTCAGCCCG 330
QY 113 ArgAspArgLysTyArgAsnGlyAspArgProAsnArgValThrSerGlyTyTrp 132
DB 331 AGGATAGAGAGTATCCAAACGGTCAAGACCTATCGGTTCCGGGTCGGGTTATGG 390
QY 133 LysAlaThrGlyAlaAspArgMetLeuArgSerGluThrSerArgProLeuGlyLys 152
DB 391 AAAGCCACCGGTACGATAAGTATCTCGACGGAG--GGAAGAAGAGTTCGTATCAAG 447
QY 153 LysThrLeuValPheTySerGlyLysAlaProLysGlyThrArgThrSerTrpLeuMet 172
DB 448 AAAGCTTTGGTGTATACATTGGAAGCTCCAAAGGAAACCAACCAATTTGGATTATG 507
QY 173 AsnGluTyArgLeu-----ProHisHisGluThrGluLysTyGlnLysAlaGluLeu 190
DB 508 CATGAGTACCGTCTCATCGAACCTCTCGTGAATGGAAGCACCAGCTTGATGG 567
QY 191 SerLeuCyAspValTyLysArgProGlyValGluAspHisProSerValProArgSer 210
DB 568 GTTTTATGTCGAATATACAAAAGCAACAAAGCGCACAAACAAAGCTTACAAATCTA 627
QY 211 LeuSerThr-----ArgHisHisAsnHisAsnSerSerThrSerSer----- 224
DB 628 ATGACGAGTGGTGGTGAATACACAAATATGGTTCGTGACATCTCTTCGTCATCAA 687
QY 225 -----ArgLeuAlaLeuArgGlnGlnHisSer 235
DB 688 TAGCAGCAGCTTCGAGCTGTTGATGATGATGATGACACAGAAAGTTGGGTTTCCGCC 747
QY 236 SerSerSerAsn-----HisSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsn 251
DB 748 GTTTCATCAACCGGCTGCTCATAGTCATAGACCGGTTTAAACCAATCATAAACCGGG 807
QY 252 -----AsnAsn 253
DB 808 TTTTCCAGGTTTAGCCAGGAGCCAGTTTGTATGGCGGAATTTGATGGACAGAACTCG 867
QY 254 LeuGluLysLeuSerThrGluTySer-----GlyAspGlySerThr 267
DB 868 GTCCCGGAACCTCGGACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
QY 268 ThrThrThrThrThr-----AsnSerAsnSerAspValThrThrAlaLeu 282
DB 928 CAGCAACAACCTGAGGGGATTCCTCGGTTTAAATAATACTCGGACTCTCG----- 978
QY 283 AlaAsnGln 285
DB 979 GCTAATCAG 987
```

RESULT 2

US-11-172-740-1242
; Sequence 1242, Application US/11172740
; Publication No. US20060057724A1
; GENERAL INFORMATION:
; APPLICANT: MASCIA, Peter
; APPLICANT: ALEXANDROV, Nikolai

APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FO
FILE OF INVENTION: PLANT CHARACTERISTICS AND PHENOTYPES
FILE REFERENCE: 2750-1602PUS2
CURRENT APPLICATION NUMBER: US/11/172.740
CURRENT FILING DATE: 2005-06-30
PRIOR APPLICATION NUMBER: 60/583,621
PRIOR FILING DATE: 2004-06-30
PRIOR APPLICATION NUMBER: 60/584,829
PRIOR FILING DATE: 2004-06-30
PRIOR APPLICATION NUMBER: 60/584,800
PRIOR FILING DATE: 2004-06-30
NUMBER OF SEQ ID NOS: 2523
SEQ ID NO 1242
LENGTH: 1606
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1606)
OTHER INFORMATION: Ceres CLONE ID no. 38214
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1606)
OTHER INFORMATION: Also known as Ceres CDNA ID no. 23373803
US-11-172-740-1242

Alignment Scores:
Pred. No.: 1,74e-41 Length: 1606
Score: 497.00 Matches: 124
Percent Similarity: 45.0% Conservative: 47
Best Local Similarity: 32.6% Mismatches: 109
Query Match: 24.4% Indels: 100
DB: 11 Gaps: 12

US-10-780-703-2 (1-379) x US-11-172-740-1242 (1-1606)

```
QY 24 IleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGluAspGluAla 43
DB 147 ATGAAAGACCCGATCTTCCTCGGTGGTCCACCACCGCCACCACTTAAC----- 194
QY 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63
DB 195 -----CTTCTCCAGCTTCGGTTTCACCTACCACGACGAG 233
QY 64 LeuIleGluPheTyLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeu 83
DB 234 CTTGTTGTTCACTACCTCAACGCAAGCAGCCTCTGCTCTTACCTGTCGCCATCATC 293
QY 84 ThrPheLeuAspLeuTyArgTyArgProThrGluLeuProAlaMetAlaAlaIleGly 103
DB 294 GCCGAGTCGATCTCTATAATTTGATCCATGGGACTTCGCGCTAAAGCATCGTTTGA 353
QY 104 GluLysGluTyTrpTyPheTyValProArgAspArgLysTyArgAsnGlyAspArgPro 123
DB 354 GAACAAGAATGGTACTCTTCTTAGTCCACGAGATCGAAAGTATCCAAACGGAGCAAGCA 413
QY 124 AsnArgValThrThrSerGlyTyTrpLysAlaThrGlyAlaAspArg---MetIleArg 142
DB 414 AACAGAGCGGCGACTTCAGGTTATTGGAAGCGCCGCTACAGATAAACCCGCTACTTGT 473
QY 143 SerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTySerGlyLysAla 162
DB 474 TCCGACGGTAAACAAAGTGGCGGTGAAGAGGCACTAGTCTTCTACAGTGGTAAACCA 533
QY 163 ProLysGlyThrArgThrSerTrpIleMetAsnGluTyArgLeu-----Pro 178
DB 534 CCRAAAGCGGTTAAAGTATGGATCATGATGATGATGATGATGATGATGATGATGATGAT 593
QY 179 His-----HisGluThrGluLysTyGlnLysAlaGlu 189
DB 594 AACATCAGACCTCTCGGCTGTGATTTTCGGCAACAAAGAAAAAACAATCACTCAGACTGTAT 653
```

QY 190 IleSerLeuCybArgValTyrLysArgProGlyValGluAspHis----- 204
DB 654 TGGGTGTTATAGTAATCTACAAGAGAACCAACGACATGTTGATAACGATAAG 713
QY 205 -----ProSerValProArgSerLeuSer----- 212
DB 714 GATCATGATATGATCGATTACATTTTCAGGAAGATTCCTCCATCTTTATCAATGGCGCT 773
QY 213 -----ThrArgHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeu 228
DB 774 GCTTCTACAGACTTCACCATCATCATATATCTCAAGATCAATGAATTTCTTC--- 830
QY 229 ArgGlnGlnHisHisSerSerSerAsnHisSerAspAsnAsnAsnAsn 248
DB 831 -----CCTGCAAAATTCCTCGGTGTTACGGGATTTCTCTGACGGTGTAAAC 881
QY 249 AsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThr 268
DB 882 ACGAGTATA-----TACGACGGCGGTGGCATGATCAAC 914
QY 269 ThrThrThrThrAnSer-----AsnSerAspValThr----- 279
DB 915 AATATTGGTACTGACTCAGTGGATCAGACAAATAACGACGACGCTGTTGTTAAATCAT 974
QY 280 -----IleAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyr 293
DB 975 GCTTCGTCGTGCTGCTATGATGATGCGGAATTTGAACGAATCTCCCGGTCCGTAT 1034
QY 293 ----- 293
DB 1035 TGGCCTGTAGCAGAGGAGGAGAGATGCATCTCCGAGCAACCGTTTCACGGTGTAGGA 1094
QY 294 -----AspThrSerAsn-----AsnThrLeuIleValSerThrArgAsnHis 307
DB 1095 GAGGAGGAGGAGATGTTTGAACATGTTCTTCGATGATGGAAGAGACACTCCACTA 1154
QY 308 GlnAspAspGluThrAlaIleValAspLeuGlnArgLeuValAsnTyrGlnIle 327
DB 1155 ATGCACAACAGGTGTTATTAGGAGATGATTTATTCAGACGACATCGTACCATTAA 1214

RESULT 3

US-11-096-568A-31996
; Sequence 31996, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Theryby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 31996
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1219)
; OTHER INFORMATION: Ceres Seq. ID no. 13591899
US-11-096-568A-31996

Alignment Scores:
Pred. No.: 2,57e-41 Length: 1219
Score: 494.00 Matches: 100
Percent Similarity: 61.3% Conservative: 33
Best Local Similarity: 46.1% Mismatches: 64
Query Match: 24.3% Indels: 20
DB: 11 Gaps: 4

US-10-780-703-2 (1-379) x US-11-096-568A-31996 (1-1219)

QY 53 ProGlyPheArgPheHisProThrGluGluLeuIleGluPheTyrLeuArgArgLys 72

DB 85 CCGGGTTTCGATTTTACCAGCGGATGAAGAGCTTATGTTCAATATCTCTGTAGAAA 144
QY 73 ValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAsp 92
DB 145 GCAGCTGGTTACGATTTCTCTCTTCAGCTCATCGCGAAATAGATCTTTACAAATTCGAT 204
QY 93 ProTyrGluLeuProAlaMetAlaIleGlyGluLysGluTyrPheTyrValPro 112
DB 205 CCATGGTTTACAAATAAAGCATTTATTGGAGAAAAGAAATGTTATTTTATGTCCT 264
QY 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTyr 132
DB 265 AGGATAGAAAATATCAACAGCGGTCAAGACCTAACCGGTTGCGGATCGGTTATTGG 324
QY 133 LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 152
DB 325 AAAGCTACGGGTACGGATAAATAATCTCGACGGAAGGACAAAGA---GTTGGTATTAAA 381
QY 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrIleMet 172
DB 382 AAAGCTTTGGTGTGTTTACATCGGAAAGCTCTTAAGGTACTAAACCAATTTGATCATG 441
QY 173 AsnGluTyrArgLeu-----ProHisHisGluThrGluLysTyrGlnLysAlaGluIle 190
DB 442 CATGAGTATGCTCTCATTTGAACCTTCTCGTAGAAACGGAAGCACTAAGTTGGATGTTGG 501
QY 191 SerLeuCybArgValTyrLysArgProGlyValGluAspHisProSerValProArgSer 210
DB 502 GTTCTATGCGAATATACAAAGCAATCAAGTGCACAAAACAAAGTTTACGATAATCGA 561
QY 211 LeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgGln 230
DB 562 ATCGCAATGCTAGAGAAATTCACCAACACGGTACTTCTGCCACG----- 606
QY 231 GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 245
DB 607 -----ACGTCGCT 654
QY 246 -----AsnAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThr 259
DB 655 GAGATCGACACACAGAAATTTCCAGTTTCTTAACCCAAACCGCATCTCGTCG 705

RESULT 4

US-11-096-568A-4700
; Sequence 4700, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Theryby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4700
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1191)
; OTHER INFORMATION: Ceres Seq. ID no. 14303479
US-11-096-568A-4700

Alignment Scores:
Pred. No.: 9,36e-41 Length: 1191
Score: 488.50 Matches: 123
Percent Similarity: 49.2% Conservative: 52
Best Local Similarity: 34.6% Mismatches: 108
Query Match: 24.0% Indels: 74
DB: 11 Gaps: 10

US-10-780-703-2 (1-379) x US-11-096-568A-4700 (1-1191)

```
QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTyrLeuArgArgLys 72
DB 115 CTGGATTTCAGATTCATCAACAGATGAAGAACTCGTCATCATCTTGTGCGCAA 174
QY 73 ValGluGlyLeuArgPheAsnValGluLeuLeuLeuThrPheLeuAspLeuTyrArgTyrAsp 92
DB 175 TGGCGTTCGCAACATATCGCGGTTCCTCAATAATCGCGAAATGATCTGACAGTACGAC 234
QY 93 ProTyrGluLeuProAlaMetAlaAlaLeuGlyGluLeuTyrPheTyrValPro 112
DB 235 CCTTGGAGCTTCCAGGAATGGCTTGTACGAGAGAGAGAGTGGTATTTTTCACGCG 294
QY 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 132
DB 295 AGGACCGCAAGTACCGCAACGGTTCGCGCGCAACCGGTTCGCGGAACCGGGTACTCG 354
QY 133 LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 152
DB 355 AAGGCAACCGGAGCGGATAACACAGTTGGTAAA-----CCGAAACCGGTGGGATCAAG 408
QY 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet 172
DB 409 AAAGCGTGTGTTTTTACGCTGGAAAGCGCCCAAGGAGAGAAACTAACTGGATCATG 468
QY 173 AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluLeuSer--- 191
DB 469 CACGAGTATCGTCTTCAGACGCTGGATCGTTCGGTTCCGCAAAAGAACAGCTTAAGGCTG 528
QY 192 -----LeuCybArgValTyrLysArgProGly----- 200
DB 529 GATGACTGGTGTCTGCCGATTTTAAACAGAAAGGTGCAATTGAAAGCAACAACTA 588
QY 201 -----ValGluAspHis----- 204
DB 589 CCACCACCGAGTGGGTTCGCCAAATTTGAATGTTCCGAAATGGAGGACGAGACCGGAG 648
QY 205 -----ProSerValProArgSerLeuSerThr----- 213
DB 649 ATTCTGCCCGCAGATCCGCGGTATACGCGCGACGCGTGGCGGATTCCTGACTTCGAG 708
QY 214 -----ArgHisHisAsnHisAsnSerSerThrSerArgLeuAla 227
DB 709 GCTTCGCACTCGGTGCGCGCGTGCACACGACGATCGAGCTGCCGACGAGGTGGTG 768
QY 228 LeuArgGlnGlnGlnHisHisSerSerSerSerAsnHisSerAspAsnAsnLeuAsnAsn 247
DB 769 TCGGCGGAGTTTCGAGCGAGGTGCAGAGCGAGCCGAGAGGGGGCAGCAACAACAAC 828
QY 248 -----AsnAsnAsnIleAsnAsnLeuGlu 255
DB 829 GAGTTTGCATATAATTACGTGGATGCCACTTCGCGGAATAATCAGATGTCGCGCTGCAG 888
QY 256 Lys-----LeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThr 271
DB 889 GATATTTTCATGTACTCTCCAAAGTCTCCAAAGTCTCCAAAGTCTCCAAAGTCTCCAAAG 947
QY 272 ThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMet 291
DB 948 AGACGAGCTGCTGTGTGTGTGTATGCATT-----ATG 983
QY 292 ProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAspAspAsp 311
DB 984 CATTTTTCAGTAGTACTCTCCACTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 1040
QY 312 GluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGlyAla 331
DB 1041 AGATCCGTGAGTGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1100
QY 332 ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIle 347
DB 1101 AGCACCAGATGAGAGAGAGAGAGATCTTAT---ATGACATGATG 1145
```

RESULT 5

```
US-11-087-099-65
; Sequence 65, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE OF INVENTION: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 65
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-087-099-65
```

```
Alignment Scores:
Pred. No.: 3,066-40 Length: 963
Score: 482.50 Matches: 100
Percent Similarity: 60.0% Conservative: 29
Best Local Similarity: 46.5% Mismatches: 53
Query Match: 23.7% Indels: 34
DB: 11 Gaps: 5
```

US-10-780-703-2 (1-379) x US-11-087-099-65 (1-963)

```
QY 24 IleGluAspAsnAspHisArgGlyGlnGluSerHisValGlnAsnGluAspGluAla 43
DB 1 ATGAGAGATACAGATTTCTTCGGTGTCTCTCCGCGCGCCCAACCAAAAC----- 48
QY 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63
DB 49 -----CTCCCTCAGAGATTCGCGTTTCATCCACACAGCAGAGAA 87
QY 64 LeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuIle 83
DB 88 CTGTGTAATTCATTACCTCAACGCAAGACAGATTCCTGTTCTTACCAGTCGCGATCATC 147
QY 84 ThrPheLeuAspLeuTyrArgTyrAspProThrGluLeuProAlaMetAlaIleGly 103
DB 148 GCGCAGCTGTGATCTTTTACAAATTTGATCCATCGGAACCTTCGCGCAAGCTTCGTTTGA 207
QY 104 GluLysGluTyrTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgPro 123
DB 208 GAACAGAGATGTGATTTTTCAGTCCAGAGATCGAAATATATCCACACGAGCTAGACCT 267
QY 124 AsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSer 143
DB 268 AACCGAGCTGCGACTTCGGTTATTGGAAGCGACTGGTACAGATAAACCGGTGATTCA 327
QY 144 Glu-----ThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
DB 328 ACCGCGGTGTGTGTAGTAGTAAAGAGTGGAGTTAAAAAGGCTCTAGTGTTTTACAGTGT 387
QY 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu----- 177
DB 388 AAACCCACCAAGAGGAGTTAAATCAGATTGATATGCAATATCGGTAACTAGTAAT 447
QY 178 ---ProHisHis-----GluThrGluLysTyrGlnLysAlaGluIle 190
DB 448 AAACCTACTCATTTTGTGACTTCGCAACAAGAAAACTCTCTCAGGCTTGATGTTG 507
QY 191 SerLeuCybArgValTyrLysArgProGlyValGluAspHisProSerValProArgSer 210
DB 508 GGTGTGTGTGTATCTACAGAAA-----AACATAGT 540
QY 211 LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArg 225
DB 541 ACAGCATCTAGACATCATCATCTCT-TCATCATATTTCATCTAGA 584
```

RESULT 6

US-11-096-568A-7863
; Sequence 7863, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7863
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1304)
; OTHER INFORMATION: Ceres Seq. ID no. 15176333
US-11-096-568A-7863

Alignment Scores:
Pred. No.: 9,05e-40 Length: 1304
Score: 479.50 Matches: 99
Percent Similarity: 59.1% Conservative: 31
Best Local Similarity: 45.0% Mismatches: 73
Query Match: 23.6% Indels: 17
DB: 11 Gaps: 5

US-10-780-703-2 (1-379) x US-11-096-568A-7863 (1-1304)

QY 18 AsnAsnAsnGluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisVal 37
DB 132 ATCTGACGAGAGAGAGATCGAGAGATGACGCGCGCGCGCGGAGGATCG----- 185
QY 38 GlnAsnGluAspGluAlaAspAspHisAspHisMetValMet----ProGlyPheArg 56
DB 186 -----GCGGCGACGACGAGGCGAGCAGGATCTGCAGCTCGCGCGCGGTTTCAGA 233
QY 57 PheHisProThrGluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLys 76
DB 234 TTCATCCGACGACGAGGAGCTGGTGATGACCTACCTGCTGCGCGCGCGCGCGCTG 293
QY 77 ArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTTPGluLeu 96
DB 294 CCATCTCCGTGCCATCTACGCGAGGTGACCTCTACAGTTTCGACCCATGCGACCTC 353
QY 97 ProAlaMetAlaAlaIleGlyLysGluTyrPheTyrValProArgAspArgLys 116
DB 354 CCGAGATGCGCTCTACGCGGAGAGAGTGGTACTTCTTCTCCCGCGCGCGCGCAAG 413
QY 117 TyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTyrPheAlaThrGly 136
DB 414 TACCCGAACGGTTCGCGCGCGAACCGGTTCGCGCGCGCGCGCTACTCGAAGGCGCACCG 473
QY 137 AlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuVal 156
DB 474 GCGACAAAGCGGTG-----GGCACCGCCCAAGCGCTTGGCCATCAAGAAGCGCTCGTC 527
QY 157 PheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrPheMetAsnGluTyrArg 176
DB 528 TTCTATGCGCGGAGGCGCGCCAGGCGGACAGACCACTGGATCATGACGAGTACCGC 587
QY 177 LeuProHisHisGluThrGluLysTyrGlnLysAlaGluIleSer----- 191
DB 588 CTCGCGACGTCGACCGCTCCGCCCGCAAGAGAACAGCAGCTCAGTTGGATGACTGGGTG 647
QY 192 LeuCysArgValTyrLysArgProGlyValGluAspHisProSerVal-----Pro 208
DB 648 CTTTCCCGGATCTACAAACAAGAGGCGCTCGGAGAGGCGCGCGCGGTGACCGGACC 707
QY 209 ArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeu 228
DB 228

Db 708 GCGAGCGCGAGCCCGCGGACGCGCGCTGCGCTGCGCGCGGAGAGAAACCGCGCTG 767

RESULT 7
US-11-096-568A-5850
; Sequence 5850, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 5850
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1396)
; OTHER INFORMATION: Ceres Seq. ID no. 14312047
US-11-096-568A-5850

Alignment Scores:
Pred. No.: 1,41e-39 Length: 1396
Score: 478.00 Matches: 116
Percent Similarity: 50.1% Conservative: 54
Best Local Similarity: 34.2% Mismatches: 108
Query Match: 23.5% Indels: 61
DB: 11 Gaps: 10

US-10-780-703-2 (1-379) x US-11-096-568A-5850 (1-1396)

QY 18 AsnAsnAsnGluLysGlyIle-----GluAspAsnAspHisArgGlyGlyGlnGlu 34
DB 70 AACAAACAACAACAAGCTTGCAAAATTTAGAAATCTTCAGTTAAAGAGGAAGAAGAAC 129
QY 35 SerHisValGlnAsnGluAspGluAlaAspAspHisAsp-HisAspMetValMet---Pr 53
DB 130 ACACAAGAGAAGAGAAGAAAGAACGCGCGCGAATGAAGGAGAAATAGAGTTGCCACC 189
QY 53 OGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyrLeuArgArgLysVa 73
DB 190 TGGTTTCAGATTTCACCCACTGATGAAGAATTGGTGAATCACTACTTGTGTAGGAAGTG 249
QY 73 lGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspPr 93
DB 250 CGCTGTCACCAATCGCGTTCCTCCGTCATCAAGAGAGTCGATTGTACAAAGTTTGATCC 309
QY 93 OTTPGluLeuProAlaMetAlaAlaIleGlyLysGluTyrPheTyrValProAr 113
DB 310 ATGGCAGCTTCCAAAATTTGGTTTTTACGCGGAGAAAGAAATGGTACTTCTTTCTCTCG 369
QY 113 GAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTyrPLy 133
DB 370 GAGCCGGAATATACCCGACGGTTCACGCGCGAAGCGCGCGCGGAGCGGTATTGGAA 429
QY 133 sAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLy 153
DB 430 AGCCACCGCGCTGTATAAACCCGATCGAAAA-----CCGAAAGCGCTTGGGATCAAGAA 483
QY 153 sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrPheMetAs 173
DB 484 AGCTCTGGTTTTTACGCGGAAAGCCCAAGGTGTGAACCAATTTGGATTATGCA 543
QY 173 nGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluIleSer----- 191
DB 544 CGAATATCGCTCCCAATTTGACCGATGTTGACCGATCTGCTCCCAAGAAAAAACACACTTGAG 603
QY 192 -----LeuCysArgValTyrLysArgProGly---ValGluAspHis-- 204
DB 604 GCTTGATGATTGGGTGTTGTGTCGATCTACAAACAAGAGGAGAGATTGAGAAATACAA 663

Qy 205 -----ProSerValProArgSerLeuSerThrArgHisHisAsnHisAs 219
 Db 664 CACAGGCGCAGGAGAGATGAATGTTGAGATGGTTTCATAGTTTGGACGACGAGAACGAGAC 723
 Qy 219 nSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerAs 239
 Db 724 GAAGCCAGAGATTCATAAGCTAGGAATGAGCAATGTGTACATGGAGACTTCGGATTGC-- 781
 Qy 239 nHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnLeuLeuGluLysLeuSerTh 259
 Db 782 -----GTGCCAAGGTTGAACAC 798
 Qy 259 rGluTyrSerGlyAspGlySerThrThrThrThrThrThrAsnSer----- 274
 Db 799 GGACTCGAGCAGCTCGGACGACGTGGTTTCGCCCGATGTCAGCTCGGAGAGGAGGTGCA 858
 Qy 275 -----AsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrAr 289
 Db 859 GAGCGACCCCAAGTGAACGATGATCTGGACCTAAGCTAGAAAC----- 904
 Qy 289 gProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAs 309
 Db 905 -----GCGTTTGATTTTCAGTTTAATTAATTG-----GA 933
 Qy 309 pAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIle 327
 Db 934 CGATAATAACCTTTCCTGGTGATGATACCTTTTGGCACTGTCAGTATCAATG 988

RESULT 8

US-10-391-414-4

; Sequence 4, Application US/10391414

; Publication No. US20050278799A1

; GENERAL INFORMATION:

; APPLICANT: SHINOZAKI, Kazuko

; APPLICANT: KATSURA, Koji

; APPLICANT: ITO, Yusuke

; TITLE OF INVENTION: Stress Induced Promoter Derived From Rice

; FILE REFERENCE: 382.1041

; CURRENT APPLICATION NUMBER: US/10/391,414

; PRIOR FILING DATE: 2003-03-18

; PRIOR APPLICATION NUMBER: JP 2002-377316

; PRIOR FILING DATE: 2002-12-26

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1423

; TYPE: DNA

; ORGANISM: Oryza sativa

US-10-391-414-4

Alignment Scores:

Pred. No.: 3,34e-39 Length: 1423

Score: 474.50 Matches: 93

Percent Similarity: 62.7% Conservative: 28

Best Local Similarity: 48.2% Mismatches: 43

Query Match: 23.3% Indels: 29

DB: 8 Gaps: 5

US-10-780-703-2 (1-379) x US-10-391-414-4 (1-1423)

Qy 27 AsnAspHisArg-----GlyGlyGlnGluSerHisValGlnAsn 39
 Db 85 GACGATCACCGGAGAGAGGAGAGAGATGAGCGCGGTCAGGACTCGAGCTGCCG--- 141
 Qy 40 GluAspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisPro 59
 Db 142 -----CCGGGGTTCGGTTCACCCG 162
 Qy 60 ThrGluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsn 79
 Db 163 ACGGACGAGGAGCTGTGTATGCACTACTCTCCCGCGCTGCGCGCCCTCCCATCGCC 222

Qy 80 ValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMet 99
 Db 223 GTCCCATCATCGCGAGATCGACCTCTACAGTTTCGATCCATCGCAGCTTCCCGGATG 282
 Qy 100 AlaAlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsn 119
 Db 283 CGCGTGTACGAGAGAGAGAGTGTCTCTCTCTCCCGCGAGACCGAGTACCCGAC 342
 Qy 120 GlyAspArgProAsnArgValThrThrSerGlyTyrTyrLysAlaThrGlyAlaAspArg 139
 Db 343 GGGTCCGCGCGCAACCCCGCGCGGTTCGGGTACTTGGAGGCGACCGCGCCGACAG 402
 Qy 140 MetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSer 159
 Db 403 CCGTGGGCTCG-----CCGAGCCGCTGGCGATCAAGAGGCGCTCGTCTCTACGCC 456
 Qy 160 GlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHis 179
 Db 457 GCGAAGGCGCCCAAGGCGAGAACCACTGGATCATGACGAGTACCGGCTCGCCGAC 516
 Qy 180 HisGluThrGluLysTyrGlnLysAlaGluIleSer-----LeuCysArg 194
 Db 517 GTCACCGCTCCGCGCCCGCAGAGAACAGACCTCAGTTGGATGATGGTCTGTGCCGG 576
 Qy 195 ValTyr---LysArgProGlyValGluAspHisProSer 206
 Db -577 ATTACAAAGAGAGGCGGCTCGAGAGAGCGCGCGCC 615

RESULT 9

US-11-096-568A-7102

; Sequence 7102, Application US/11096568A

; Publication No. US20060048240A1

; GENERAL INFORMATION:

; APPLICANT: Alexandrov, Nikolai et al.

; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide

; FILE REFERENCE: 2750-1592PUS2

; CURRENT APPLICATION NUMBER: US/11/096,568A

; CURRENT FILING DATE: 2005-04-01

; NUMBER OF SEQ ID NOS: 34471

; SEQ ID NO 7102

; LENGTH: 1395

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(1395)

; OTHER INFORMATION: Ceres Seq. ID no. 15170403

US-11-096-568A-7102

Alignment Scores:

Pred. No.: 4,14e-39 Length: 1395

Score: 473.50 Matches: 115

Percent Similarity: 46.2% Conservative: 44

Best Local Similarity: 33.4% Mismatches: 79

Query Match: 23.3% Indels: 107

DB: 11 Gaps: 9

US-10-780-703-2 (1-379) x US-11-096-568A-7102 (1-1395)

Qy 53 ProGlyPheArgPheHisProThrGluGluLeuIleGluPheTyrLeuArgArgLys 72
 Db 104 CCAGGCTTCAGATTCATCCAAACGAGCAGGAGCTGGTGTGCATCTCTCGCGTAAA 163
 Qy 73 ValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAsp 92
 Db 164 TGGCATCGCAGCTATCGCGTTCCTCCATCATCGCGAAATCGACCTCTACAAATACGAC 223
 Qy 93 ProThrGluLeuProAlaMetAlaAlaIleGlyGluLysGluTyrPheTyrValPro 112
 Db 224 CCCTGGGACCTACCGGATTCGCTCCCTACCGAGAGAAAGAGTGGTACTCTTTTCACCA 283
 Qy 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 132

```
Db 284 CGGACCGGAAATACCCGACCGTTCCAGGCGGACCGCGCGGGGACCGGTTACTGG 343
Qy 133 LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 152
Db 344 AAGCAACCGCGGGGATGAAACCCATTCGCCAC-----CCCAACCGTTGGGATATAA 397
Qy 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet 172
Db 398 AAGCTTTGGTGTGTTTACCGAGGAAAGCTCCGAAAGGGGACAGAGCAATGGATCATG 457
Qy 173 AsnGluTyrArgLeuProHisGluThrGluLysTyrGlnLysAlaGluSer---- 191
Db 458 CACGAGTATCGTTCGCCGATGTAGATCGCTCCGTTCCGTAACAAAGAGGCGACGCTTAAGGTTA 517
Qy 192 -----LeuCysArgValTyrLysArgProGly----- 200
Db 518 GATGATTGGGTGCTTTGCGGTATTACAAAGAGAGGCGACGATCGAAGAGTTCCAAACCA 577
Qy 201 -----ValGluAspHis----- 204
Db 578 AGCAGCGATGTGTGTTAGCGGAAATGAATCATCGGAGATCGAAGACAGGAGCGG 637
Qy 205 -----ProSerValPro-ArgSerLeuSerTh 213
Db 638 GAGATTCTGAAAGCGGAGGAGTGTCTTCTGCGCGCGGTTCGCCGCCCAAGGGAAG 697
Qy 213 rArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHis 233
Db 698 GCGCGGTGAAGAGGATTTACATGTACTTTCACCGGT---CGGATTCAATCCCGAAGGTG 754
Qy 233 sHisSerSerSer----- 238
Db 755 -CACACGACTCGAGTGTTCGGAGCAGCTGGTATCGCGGAAATTCGGACGCGAGGTGCA 813
Qy 239 -----AsnHisSerAs 242
Db 814 GAGCGAGCAAGTGAAGGAGTGGGAGAAAGCTCGAGTTCCGTTTAAATACGTGA 873
Qy 242 pAsnAsnLeuAsnAsnAsnAsnAsnHisAsnAsnLeuGluLysLeuSerThrGluTyrSe 262
Db 874 TGCCACTCTGAACACACAGCAAGCTTC----- 901
Qy 262 rGlyAspGlySerThrThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLe 282
Db 902 -----ACGACGCAATTCAGGCGCAATTAATCAGATGATGTCCCGCT 942
Qy 282 uAlaAsnGlnAsnIleTyrArgPro---MetProTyrAspThrSerAsnAsnThr----- 299
Db 943 GCAGGACATGTTGATGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002
Qy 300 -----LeuIleValSerThrArgAsnHis 307
Db 1003 CATTCGCAATTCGCGGTAGCAGCAGACTCAAGGTCTATGATGTGTGCTGTCGACCA 1062
Qy 307 sGlnAspAsp 310
Db 1063 CACAGACGAT 1072
```

RESULT 10

```
US-11-096-568A-26116
; Sequence 26116, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26116
; LENGTH: 1189
```

```
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1189)
; OTHER INFORMATION: Ceres Seq. ID no. 13498621
US-11-096-568A-26116
```

```
Alignment Scores:
Pred. No.: 6,23e-39 Length: 1189
Score: 471.00 Matches: 103
Percent Similarity: 54.6% Conservative: 57
Best Local Similarity: 35.2% Mismatches: 85
Query Match: 23.2% Indels: 48
DB: 11 Gaps: 8
```

US-10-780-703-2 (1-379) x US-11-096-568A-26116 (1-1189)

```
Qy 43 AlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGlu 62
Db 58 GCGCGGAGCAGCAGCATCAGCTGGATCTCGCGGGTTCGGTTCCACCCGACGAGGAG 117
Qy 63 GluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeu 82
Db 118 GAGCTGTGGAGTTCTACCTGAAGCAGCGCGCGCGGAGGAGCTCAAGTTCCGATC 177
Qy 83 IleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAla---Ala 101
Db 178 ATCCCAAGGTGACCTGTACCGGACGACCCCTGGGAGCTGCGAGCTGGCGCGGCTC 237
Qy 102 IleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArg----- 118
Db 238 GTTGGCGAGCGGAGTGTACTTCTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 297
Qy 119 -----AsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAla 134
Db 298 CAGCAGCGCGGCTCCGCGCGCGCTTAGCGCACACGAGCGCGCGGTTCTGGAAGGCC 357
Qy 135 ThrGlyAlaAspArgMetIleArg-----SerGluThrSerArgProIleGlyLeuLys 152
Db 358 ACGGGGTCGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCATCGGCTCAAG 417
Qy 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet 172
Db 418 AAGAGCTGCTCTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATG 477
Qy 173 AsnGluTyrArgLeuProHis-----HisGluThrGluLysTyrGlnLysAla 188
Db 478 AACGAGTACCGCTCCCGCGAGCTGCGCGAGCTAGCAGCAGCGCGCGCGCGCGCGCGGTG 537
Qy 189 GluIleSerLeuCysArgValTyrLys-----Arg 198
Db 538 GAGCTGTGTGTGCAAGGTCTACCGCAAGCGCGTGTGCTCAAGGAGCTGGAGCAGCGG 597
Qy 199 ProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisAsnHis 218
Db 598 GTGGCCATGGAGGAG-----CTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 648
Qy 219 -----AsnSerSerThrSerSerArgLeuAlaLeu----- 228
Db 649 AGCACCGGCTCCCGCGCGAGCTCATGCTGCTGCGGAGATCGTGTGTTGCTGCGCGTGG 708
Qy 229 -----ArgGlnGlnGlnHisSerSerSer 237
Db 709 GCGCTCTCGTCTCACCGTCAAGTGCATGAGCATGAGAGAGAGAGAGCGCGCGCGAAGANT 768
Qy 238 SerAsnHisSerAspAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnLeuGluLysLeu 257
Db 769 CAGCAGCAGCAGGACCCCTCTCTGACGACGCTGCGGAGCGCGCGCGCGCGCGCGCGTGG 828
Qy 258 SerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThr 277
Db 829 TCTGCTTACTACGCGCGCTCCTCAACTTCTTAACCAAAATCCACAAATCCACAGCTTT 888
```


Db 775 TTATCTGAGAAATCATACCAACCCACCCAGGGGTGGATCAAGTCTGGATTCAACACT 834
 Qy 252 AsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThr 271
 Db 835 GGGAAATTGGATCAGCAATGCTCTCCAATCAATAACAACACCCCTACCAATAATTACATG 894
 Qy 272 ThrAsnSerAspValThrIleAlaLeuAlaAsn---GlnAsnIleTyrArgPro 290
 Db 895 TTTCAGAGAATCCCAATTGAACCTTCATCTCCACATCGGAACATGATGAGGCCA 954

RESULT 13.

US-11-096-568A-24275
 ; Sequence 24275, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; TITLE OF INVENTION: Therby
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 24275
 ; LENGTH: 1833
 ; TYPE: DNA
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1833)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12424091
 US-11-096-568A-24275

Alignment Scores:
 Pred. No.: 2,32e-36 Length: 1833
 Score: 448.50 Matches: 112
 Percent Similarity: 46.2% Conservative: 46
 Best Local Similarity: 32.7% Mismatches: 105
 Query Match: 22.1% Indels: 80
 DB: 11 Gaps: 9

US-10-780-703-2 (1-379) x US-11-096-568A-24275 (1-1833)

Qy 51 ValMetPro-----GlyPheArgPheHisProThrGluGluGluLeuIleGlu 66
 Db 248 GTATGTCACCACTCCCGCGCGGTTCGGTTCACCCACGAGGAGGACTCATGGT 307
 Qy 67 PheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuIleThrPheLeu 86
 Db 308 CACTACCTCATGAGCAGCGCGCTCCATGCGGTGCGCTCCCATCATCGCGGAGTC 367
 Qy 87 AspLeuTyrArgTyrAspProThrGluLeuProAlaMetAlaIleGlyGluLysGlu 106
 Db 368 AACATCTACCAGTGCACACCCCTGGGATCTCCCTGCCAAGGCATTTGTTCCGCGACAGGAG 427
 Qy 107 TrpTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgVal 126
 Db 428 TGGTTTTCTTCACGCCCCCGGAGCGCAAGTACCCCAACGCGCGCCCAACCGCGCC 487
 Qy 127 ThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSer 146
 Db 488 GCCGGTCCGGTACTGAGAGCCACCGGCACCCAGGCGCATCTCTGCG---TCGTCC 544
 Qy 147 ArgPro-----IleGlyLeuLysLysThrLeuValPhe 157
 Db 545 ACGCCACAGCAGCCACGGCGGCCAACATCTGCTGCGCGTCAAGAAGCGCTCGTCTTC 604
 Qy 158 TyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu 177
 Db 605 TACGGCGGCGCGCGCCCAAGGGCACCAGAGGACTGGATCATGACGAGTACCGGCTC 664
 Qy 178 Pro-----HisIleGluThrGluLysTyrGlnLysAlaGluIle 190
 Db 665 TCGGGCGCGGAGCAGGACTGCAAGGGAGCCAGGCGCAGGCGCAGAGTCTTCTCTCTCG 724

Qy 191 Ser-----LeuCysArgValTyrLysArgProGlyVal 201
 Db 725 TCTCTCTCCATGAGGCTGGAGCAGTGGGTCTGTGCGAGGATCCACAAGAAG----- 775
 Qy 202 GluAspHisProSerValProArgSerLeuSerThrArgHisAsnHisAsnSerSer 221
 Db 776 -----AGC 778
 Qy 222 ThrSerSerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSerSerAsnHisSer 241
 Db 779 AACGATTTCCAGTTGTCTGCTCGGAGCATGAGCAGCAGCAGGAGGAGCGCGCGCG 838
 Qy 242 AspAsnAsnLeuAsnAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyr 261
 Db 839 GCGCGCTCACGAGCAGCTGGAGGAGTACTAGAGCTCTCTGCTGCTCGACACAGCAGC 898
 Qy 262 SerGlyAspGlySerThrThrThrThrThrAsnSerAsnSerAspValThrIleAla 281
 Db 899 AGCAGCGACACACCACCACCACCACCACCACCACCACCACCACCACCAGCTGCG 958
 Qy 282 LeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeuIle 301
 Db 959 AC----- 960
 Qy 302 ValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleVal----- 316
 Db 961 ---TCATCAGCATCTCAGACGATGATGATGCTGAGCAAGTCTGCTGCTCACCAGCT 1017
 Qy 317 AspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGlyAlaThrThrLeuMetPro 336
 Db 1018 CCTCAGCAGCATCGACTACGCGCGCTCTCTCGAGATGCTCTCTCGACGCGGCCA 1077
 Qy 337 GlnThrGln-----AlaAlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsn 354
 Db 1078 GCCCCAGCTGATGAGCGCGCGCCACACATGATGCTACTACCCCGCGCGCTCGCGGAC 1137
 Qy 355 AsnAla 356
 Db 1138 GATGCA 1143

RESULT 14

US-11-096-568A-22731
 ; Sequence 22731, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
 ; TITLE OF INVENTION: Therby
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 22731
 ; LENGTH: 1393
 ; TYPE: DNA
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1393)
 ; OTHER INFORMATION: Ceres Seq. ID no. 12409839
 US-11-096-568A-22731

Alignment Scores:
 Pred. No.: 2,12e-36 Length: 1393
 Score: 447.50 Matches: 83
 Percent Similarity: 69.5% Conservative: 24
 Best Local Similarity: 53.9% Mismatches: 39
 Query Match: 22.0% Indels: 8
 DB: 11 Gaps: 2
 US-10-780-703-2 (1-379) x US-11-096-568A-22731 (1-1393)

QY 53 ProGlyPheArg-PheHisProThrGluGluGluLeuIleGluPheTyrLeuArgArgly 72
Db 112 CCGGGGTTCCGGTTTCCACCGGACGAGGAGTGGTGCATCTCTGCGCGG 171
QY 72 sValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAs 92
Db 172 CTGCGCGCGCTGCCCATCGCGTCCCATCGCGGAGATCGACCTCTCAAGTTTGA 231
QY 92 pProTrpGluLeuProAlaMetAlaIleGlyGluLysGluTyrPheTyrValPr 112
Db 232 CCCATGGCAGCTCCCAAGAGTGGCTGTACGCGGAGAGGAGTGTACTCTTCTCCCC 291
QY 112 oArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTr 132
Db 292 GCGGACCCGACGATACCCGACGGGTCCAGGCCCAACCGCGCGCGGTGGGTACTG 351
QY 132 pLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 152
Db 352 GAAGGCCACCGCGCTGACAAAGCCCTG-----GGCACGCCCAAGCGCTGCCCATCAA 405
QY 152 sLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleWe 172
Db 406 GAAGCGCTCGTCTTCTACGCGGCAAGGCGGCCCAAGGGCGAGAGCACTGGATCAT 465
QY 172 tAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluIleSer-- 191
Db 466 GCACGAGTACCGCTCGCGACGTCGCGCGCTCGCGCGCGCAGAGAGAGCGCTCAGGTT 525
QY 192 -----LeuCyArgValTyrLysArgProGly 200
Db 526 CGATGACTGGTCTCTGCGCATCTTACCAAGAGAGGCG 565

RESULT 15

US-11-096-568A-30488
; Sequence 30488, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30488
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(897)
; OTHER INFORMATION: Ceres Seq. ID no. 4961610
US-11-096-568A-30488

Alignment Scores:
Pred. No.: 2,57e-36 Length: 897
Score: 444.50 Matches: 92
Percent Similarity: 54.8% Conservative: 51
Best Local Similarity: 35.2% Mismatches: 93
Query Match: 21.9% Indels: 25
DB: 11 Gaps: 7

US-10-780-703-2 (1-379) x US-11-096-568A-30488 (1-897)

QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyrLeuArgArgly 72
Db 13 CCAGGATTCAGGTTTCTCCACTGATGAGAACTGGTGGCTTACTTCTGTAGAGAG 72
QY 73 ValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAsp 92
Db 73 GTCAACGGCCCAAGCCATTGAGCTCGAGATCATCCAGAGAGTTGATCTTTATAAATCGGAG 132
QY 93 ProTrpGluLeuProAlaMetAlaIle-----GlyGluLysGluTrpTyrPheTyr 110

Db 133 CCATGGGACTTGCCTGAAAAGTCATTTTGGCGGAAACGACATGATGGTACTTTTAC 192
QY 111 ValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGly 130
Db 193 AGCACAGGGGATAAGAGTATATCCAAATGGCTCTAGGACGAACCGTGCACCGCGGGT 252
QY 131 TyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGly 150
Db 253 TACTGGAAGGCCACCGGAAAGATCTGACGTAGTAAATCAAGAAGATGAAG--ATGGGA 309
QY 151 LeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrp 170
Db 310 ATGAAGAAGACACTGGTTTATTATAGAGGAAGGGCTCTCATGGCTTCTGTTACTAATTCG 369
QY 171 IleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluIle 190
Db 370 GTCATGCAATGATATCGTCTCACGCACGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 426
QY 191 SerLeuCyArgValTyrLysArgProGlyValGluAspHisProSerValProArgSer 210
Db 427 GCATTGTCCGAGTGTTTAAGAAG-----AACATACAAATTCCAAAGAGA 471
QY 211 LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgGln 230
Db 472 AAAGGGGAAGAAGAAGACAGCAGACAGACTAGT-----GTAGGAAAAGAGAG 525
QY 231 GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 250
Db 526 GAAGAAAGAAAGGAG 585
QY 251 IleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThr 270
Db 586 TTGAAAAGA-----GCATCCCGGAGAGACATCT 612
QY 271 ThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgPro 290
Db 613 TCATCAGAGCTAACTCAAGGGTCTCTTTAGACGAAGCAACAGCTCATCCATTTTGTCT 672
QY 291 MetProTyrAspThrSer-----AsnAsnThrLeuIleValSerThrArgAsnHis 307
Db 673 CTTCAATTTCTCATCAT 732
QY 308 Gln 308
Db 733 CAG 735

Search completed: April 6, 2006, 10:34:43
Job time : 588 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 09:58:27 ; Search time 239 Seconds
(without alignments)
2818.811 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAIVSSTTSIIPMSNQVNNN.....MNPVDPDGNRDHYTNIPFK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlh
-Q=/abss/ABSSWEB.spool/US10780703/runat_06042006_101342_6427/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss05h -USER=US10780703 @CGN_1_193 @runat_06042006_101342_6427
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA:*

1:	/cgn2_6/ptodata/1/ina/1/COMB.seq:*
2:	/cgn2_6/ptodata/1/ina/5/COMB.seq:*
3:	/cgn2_6/ptodata/1/ina/6A/COMB.seq:*
4:	/cgn2_6/ptodata/1/ina/6B/COMB.seq:*
5:	/cgn2_6/ptodata/1/ina/H/COMB.seq:*
6:	/cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
7:	/cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8:	/cgn2_6/ptodata/1/ina/RE.COMB.seq:*
9:	/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	509	25.0	1209	3	US-09-614-408-6
2	498.5	24.5	1211	3	US-09-614-408-1
3	498.5	24.5	1211	3	US-09-614-981-1
4	494	24.3	1212	3	US-09-533-029-3
5	480	23.6	1409	3	US-09-533-029-37
6	462	22.7	1216	3	US-09-533-029-121
7	455	22.4	914	3	US-09-533-029-23
8	403	19.8	1287	3	US-09-889-926-1
9	357	17.6	913	3	US-09-533-029-73

c	10	330	16.2	1332	3	US-09-533-029-115	Sequence 115, App
	11	293	14.4	2680	3	US-09-614-408-7	Sequence 7, Appli
	12	239.5	11.8	1339	3	US-09-533-029-99	Sequence 99, Appl
	13	131	6.4	1853	3	US-09-248-796A-336	Sequence 336, App
	14	128	6.3	2094	3	US-09-248-796A-1192	Sequence 1192, A
	15	125.5	6.2	934	3	US-09-270-767-12547	Sequence 12547, A
c	16	122.5	6.2	1086	2	US-08-415-751-47	Sequence 47, Appl
	17	122.5	6.0	1107	3	US-09-614-221A-398	Sequence 398, App
	18	122.5	6.0	1974	3	US-09-248-796A-11107	Sequence 11107, A
	19	122	6.0	2001	3	US-09-487-558B-35	Sequence 35, Appl
	20	121.5	6.0	2190	3	US-09-625-188-19	Sequence 19, Appl
	21	121	5.9	750	3	US-09-248-796A-9460	Sequence 9460, Ap
	22	120	5.9	1431	3	US-09-248-796A-5089	Sequence 5089, Ap
	23	118.5	5.8	2115	3	US-09-614-221A-111	Sequence 111, App
	24	118.5	5.8	2115	3	US-09-487-558B-217	Sequence 217, App
	25	116	5.7	285	3	US-09-313-294A-2129	Sequence 2129, Ap
	26	115.5	5.7	1020	3	US-09-248-796A-1997	Sequence 1997, Ap
	27	115.5	5.7	1230025	3	US-09-198-452A-1	Sequence 1, Appli
	28	115.5	5.7	1230230	3	US-09-438-185A-1	Sequence 1, Appli
	29	114.5	5.6	789	3	US-09-248-796A-11671	Sequence 11671, A
	30	114.5	5.6	2574	3	US-09-248-796A-6419	Sequence 6419, Ap
	31	113.5	5.6	723	3	US-09-248-796A-13410	Sequence 13410, A
	32	113.5	5.6	867	3	US-09-216-393B-340	Sequence 340, App
	33	113.5	5.6	867	3	US-09-216-393B-342	Sequence 342, App
	34	113.5	5.6	1397	3	US-09-216-393B-343	Sequence 343, App
c	35	113.5	5.6	1397	3	US-09-216-393B-345	Sequence 345, App
	36	113.5	5.6	1814	3	US-09-892-360-1	Sequence 1, Appli
	37	113.5	5.6	2710	3	US-09-949-016-1497	Sequence 1497, Ap
	38	113.5	5.6	2730	3	US-09-799-451-432	Sequence 432, App
	39	113.5	5.6	2730	3	US-09-949-016-1130	Sequence 1130, Ap
	40	112.5	5.5	3456	3	US-09-487-558B-133	Sequence 133, App
	41	111.5	5.5	65848	3	US-09-949-016-13285	Sequence 13285, A
c	42	111	5.5	1225	3	US-09-270-767-11171	Sequence 11171, A
	43	111	5.5	1800	3	US-09-248-796A-3127	Sequence 3127, Ap
	44	111	5.5	2277	3	US-09-487-558B-223	Sequence 223, App
	45	110	5.4	2733	3	US-08-997-685A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-614-408-6
; Sequence 6, Application US/09614408
; Patent No. 6762347
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: VREBALOV, JULIA
; APPLICANT: NOENSIE, FREDERICK
; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
; FILE REFERENCE: TANK:213
; CURRENT APPLICATION NUMBER: US/09/614,408
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,357
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Tomato
US-09-614-408-6

Alignment Scores:
Pred. No.: 6.16e-48
Score: 509.00
Percent Similarity: 52.6%
Best Local Similarity: 35.0%
Query Match: 25.0%
DB: 3
Matches: 1209
Conservative: 57
Mismatches: 89
Indels: 64
Gaps: 10

US-10-780-703-2 (1-379) x US-09-614-408-6 (1-1209)

```
QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuLeuLeuPheTyrLeuArgArgLys 72
Db 87 CCGGGGTTTCGATTCCACCGCGGAGGAACTCATCTGCACCTACCTCAAAAACGA 146
QY 73 ValGluGlyLysArgPheAsnValGluLeuLeuLeuThrPheLeuAspLeuTyrArgTyrAsp 92
Db 147 GTCGCGCGCTCCGATTCGCGTGGATATTATTGGTGAATGATCTTTATAAGTTTGTAT 206
QY 93 ProTrpGluLeuProAlaMetAlaAlaLeuGlyGlyLysGluTrpTyrPheTyrValPro 112
Db 207 CCATGGGAATCCCTCGTGAAGCAATATTCCGGAGACCAAGATGGTCTCTTTTATGATCCA 266
QY 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 132
Db 267 AGAGATAGAAATATCTTAACGGGCGGAGCCCAATCCGGCTGCACATCGGGTATTG 326
QY 133 LysAlaThrGlyAlaAspArg---MetIleArgSerGluThrSerArgProIleGlyLeu 151
Db 327 AAGGCTACCGGAACCGACAAAGCGGTTTACTTCCGGTGAACACAAAGAGTTGGGTA 386
QY 152 LysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIle 171
Db 387 AAAAGCGCTCGTTTTTACGCGCGTAAACCAACCAAAAGGGGTAAACAACTAATTGGATC 446
QY 172 MetAsnGluTyrArgLeuProHisHisGluThrGluLys----- 184
Db 447 ATGCATGAATACAGAGTTGTAGAAATTAACAAATAACACGACCTTGGTGTGATAT 506
QY 185 -----TyrGlnLysAlaGluIleSer-----LeuCysArgValTyr 196
Db 507 ATTGTTGCCAACAAAGAGGATCTTTGAGGCTAGATGATTGGGTTTATGTCGAATTTAC 566
QY 197 LysArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHis 216
Db 567 AAGAAGATAACACACAA-----AGGTCCATAGATGATTGTCATGAT 608
QY 217 -----AsnHisAsnSerSerThrSerArgLeuAlaLeuArgGlnGln 231
Db 609 ATGTTGGGATCGATACCAAAATGTACCAAAATTCATATTAACAGGAATAAGCCTTCA 668
QY 232 GlnHisHisSer-----SerSerAsnHisSerAspAsnAsnLeuAsn 247
Db 669 AACTATGTTGACATATTGTCGAAATGAATCGAATATGTACGATGGAATTTATGAATAAC 728
QY 248 -----AsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu----- 260
Db 729 ACCAAGCATATTATCAACAATAATAATAGATCCATTCACAAATATCGTCAAGAGAACG 788
QY 261 -----TyrSerGlyAspGlySerThrThrThrThrThrThrAsnSer 274
Db 789 ATGCATGGAGGTTGTTGTAATAACGACGAGCAACCAACCAACCAACT----- 842
QY 275 AsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAsp 294
Db 843 -----ATTGATAGAAC 854
QY 295 ThrSerAsnAsnThrLeuLeuValSerThrArgAsnHisGlnAspAspGluThrAla 314
Db 855 CATTCTCCAAATACAAAAGGTTCTTGTGTGAGAACCAACGAGCAGCATGCACTTAAC----- 911
QY 315 IleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGlyAlaThrThrLeu 334
Db 912 ---ATGAATAATATTTCCGGAATTAACAATCATGAACAAAGTAGCTCCATTTGCCAATTTTC 968
QY 335 MetProGln 337
Db 969 CTGAGCCAG 977
```

RESULT 2

US-09-614-408-1
; Sequence 1, Application US/09614408
; Patent No. 6762347
; GENERAL INFORMATION:

; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: VREBALOV, JULIA
; APPLICANT: NOENSIE, FREDERICK
; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
; FILE REFERENCE: TANK:213
; CURRENT APPLICATION NUMBER: US/09/614,408
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,357
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1211
; TYPE: DNA
; ORGANISM: Tomato
; US-09-614-408-1

Alignment Scores:

Pred. No.: 9,72e-47 Length: 1211
Score: 498.50 Matches: 116
Percent Similarity: 51.9% Conservative: 61
Best Local Similarity: 34.0% Mismatches: 98
Query Match: 24.5% Indels: 66
DB: 3 Gaps: 11

US-10-780-703-2 (1-379) x US-09-614-408-1 (1-1211)

```
QY 37 ValGlnAsnGluAspGluAlaAspAspHisAspHisAsp---MetValMetProGlyPhe 55
Db 36 ATGGAAGTACGAGTTCATCAACCGGACACGTCATCAGCCTCAACTCCCGGGGTTT 95
QY 56 ArgPheHisProThrGluGluLeuLeuLeuGluPheTyrLeuArgArgLysValGluGly 75
Db 96 CGATTCCACCGGACGAGCAAGAACTCATGCTCACTACCTCAAAACCAACGCTGCCGCGC 155
QY 76 LysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAspProTrpGlu 95
Db 156 GCTCCGATTCGCGTGGATATTATTCGTAATTTGATCTTTATAAGTTTCATCCATGGGAA 215
QY 96 LeuProAlaMetAlaAlaLeuGlyGlyLysGluTrpTyrPheTyrValProArgAspArg 115
Db 216 CTCCTCGCTAAGGCAATATTCGAGAGCAAGAAATGGTCTCTTTTATGTCCAAGAGATAGA 275
QY 116 LysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThr 135
Db 276 AAATATCTTAACGGGCGGAGGCAAAATCGGGCTGCAACATCGGGTTATTGGAGGCTACC 335
QY 136 GlyAlaAspArg---MetIleArgSerGluThrSerArgProIleGlyLeuLysLysThr 154
Db 336 GGAACCGCAAGCCGGTTTTTACTTCCGGTGAACACAAAGAGTTGGGTAAAAAGGCG 395
QY 155 LeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGlu 174
Db 396 CTCGTTTTTACGCGGTAAACCAACCAAAAGGGGTAAACAACTAATTGGATCATGCATGAA 455
QY 175 TyrArgLeuProHisHisGluThrGluLys----- 184
Db 456 TACAGAGTTGTAGAAAATAAAACAAATAACAGCCACTGGTTGTGATATAATTGTTGCC 515
QY 185 TyrGlnLysAlaGluIleSer-----LeuCysArgValTyrLysArgPro 199
Db 516 RACAAAAAAGGATCTTTGAGGCTAGATGATTGGGTTTATGTCGATTTTACAAGAAGAT 575
QY 200 GlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHis----- 216
Db 576 AACACACAA-----AGGTCCATAGATGATTGTCATGATGATGTTGGGA 617
QY 217 -----AsnHisAsnSerSerThrSerArgLeuAlaLeuArgGlnGlnHisHis 234
Db 618 TCGATACCACAAAATGTACCAAAATTTCAATATTACAGGAATAAGCCCTTCAAACTATGTT 677
QY 235 Ser-----SerSerSerAsnHisSerAspAsnAsnLeuAsn----- 247
```


APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Reuber, Lynne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3

LENGTH: 1212
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G759
US-09-533-029-3

Alignment Scores:
Pred. No.: 3,17e-46 Length: 1212
Score: 494.00 Matches: 100
Percent Similarity: 61.3% Conservative: 33
Best Local Similarity: 46.1% Mismatches: 64
Query Match: 24.3% Indels: 20
DB: 3 Gaps: 4

US-10-780-703-2 (1-379) x US-09-533-029-3 (1-1212)

QY	53	ProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeuArgArgLys	72
DB	53	CGGGTTCCGATTTTACCGCGGATGAGAGCTTATGTTCAATATCTCTGTAGAAA	112
QY	73	ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp	92
DB	113	GCAGCTGGTTACGATTTCTCTCTCAGCTCATCGCGAAATAGATCTTTACAAATTCGAT	172
QY	93	ProThrGluLeuProAlaMetAlaAlaLeuGlyGluLeuTyrPheTyrValPro	112
DB	173	CCATGGGAGCTTCCGTTAGCTTGTATGGTGAAGAGGATGCTCTCTCTCTCCC	232
QY	113	ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSerGlyTyrTrp	132
DB	233	AGGATAGAAAATATCCAAACGGTCAAGACCTTAACCGGTTCCGGATCGGTTATTCG	292
QY	133	LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys	152
DB	293	AAAGCTACGGTACGGTAAATAATCTCGACGGAAGGACAAAGA--GTTGGTATTAA	349
QY	153	LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet	172
DB	350	AAAGCTTGGTGTATACATCGGAAAGCTCTTAAAGGTACTAAACCAATTCGATCATG	409
QY	173	AsnGluTyrArgLeu-----ProHisHisGluThrGluLysTyrGlnLysAlaGluIle	190
DB	410	CATGAGTATCGTCTCATTTGAACTTCTCGTAGAAACGGAAGCACTAAAGTTGATGG	469
QY	191	SerLeuCyArgValTyrLysArgProGlyValGluAspHisProSerValProArgSer	210
DB	470	GTTCTATGTCGAATATACAAAGAGCAATCAAGTGCAAAACAAAGTTTACGATATGA	529
QY	211	LeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgGln	230
DB	530	ATCGCGAATGTAGAGAAATTCAGCAACAACGGTACTTCTGTCACG-----574	
QY	231	GlnGlnHisHisSerSerSerAsnHisSerAspAsnLeu-----245	
DB	575	-----ACGTGCTTCTTCTCACTTTGAAAGACGTTCTTGATTCGTTTCATCAA	622
QY	246	-----AsnAsnAsnAsnAlaLeuAsnAsnLeuGluLysLeuSerThr	259
DB	623	GAGATCGAACACAGAAATTTCCAGTTTCTTAAACCCAAACCGCATCTGTCG	673

RESULT 5

US-09-533-029-37
Sequence 37, Application US/09533029
Patent No. 666446
GENERAL INFORMATION:

APPLICANT: Heard, Jacqueline
APPLICANT: Broun, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Keddle, James
APPLICANT: Pineda, Omalra
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Reuber, Lynne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES

FILE REFERENCE: MBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 37
LENGTH: 1409
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G502
US-09-533-029-37

Alignment Scores:
Pred. No.: 1.6e-44 Length: 1409
Score: 480.00 Matches: 112
Percent Similarity: 51.5% Conservative: 39
Best Local Similarity: 38.2% Mismatches: 79
Query Match: 23.6% Indels: 64
DB: 3 Gaps: 7

US-10-780-703-2 (1-379) x US-09-533-029-37 (1-1409)

QY	53	ProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeuArgArgLys	72
DB	248	CCAGGTTTCCGATTTTACCCCTACCGATGAAGAGCTTGTGCATGCATCTCTCTCCGCAA	307
QY	73	ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp	92
DB	308	TGTGCTCTCATGTCATCGCGTTCCGATCATCGCTGAGATCGATCTCTCAAAATACGAT	367
QY	93	ProThrGluLeuProAlaMetAlaAlaLeuGlyGluLysGluTyrPheTyrValPro	112
DB	368	CCATGGGAGCTTCCGTTAGCTTGTATGGTGAAGAGGATGCTCTCTCTCTCCC	427
QY	113	ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSerGlyTyrTrp	132
DB	428	AGGACAGAAAAATATCCCAACGGTTCGGCTCTTAACCGCTCGCTGTTCTGTTACTGG	487
QY	133	LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys	152
DB	488	AAAGCTACCGGAGCTGATAAACCCGATC-----GGACTACTAAACCGGATTAAG	541
QY	153	LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet	172
DB	542	AAAGCTCTTGTCTTCTACCGCGCAAGAGCTCCAAAGGAGGAGAAAAACCAATTCGATCG	601
QY	173	AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAla-----188	
DB	602	CACGAGTACCGTCTCGCGGACGTTCCACCGTTCGCGAAGAGAAATAGTCTCAGG	661

```
QY 189 -----GluIleSerLeuCyArGValTyrLysArgProGly----- 200
Db 662 CTGGATGATTGGGTCTCTCGCGGATTTCACAAAGAGGAGCTACCGAGAGCGCGGA 721
QY 201 -----ValGluAepHisProSerValPro-Ar 209
Db 722 CCACCGCCTCGGTGTTTACGGCGACGAAATCATGGAGGAGAACCGCAAGGTGACGGAG 781
QY 209 qSerLeuSerThrArgHisAsnAsn---SerSerThrSerSerArgLeuAlaLe 228
Db 782 ATGGTTATGCTCCCGCGCCCAACAGCAAGTGGTTCGGTATTTCGACACGTCGGAT 841
QY 228 uArgGlnGlnHisSerSerSerSerSerSerSerSerSerSerSerSerSerSer 246
Db 842 TCGGTGCGGAAGCTG-CATACTACGGATTTCGAGTTGCTCGGAGCAGGTGGTGTGCGCGGA 900
QY 247 -----AsnAsnAs 249
Db 901 GTTCACGAGCGAGGTTTCAGAGCGAGCCCAAGTGGAAGATTGGTTCGCCCGCTAAGTAATGA 960
QY 249 nAsnLeuAsnLeuGlu-----LysLeuSerThrGluTy 261
Db 961 CAATAACATACCTCTGTTTGGGTTTAAATCATGATGCCACCGTGGATTAACGGCTT 1020
QY 261 rSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 281
Db 1021 TGGAGGAGGAGGAGT-----AGTAATCAGATGTTTCC 1053
QY 281 aLeuAlaAsnGlnAsnIleTyrArgProMetProTyr 293
Db 1054 GCTACAGGATATGTTTCATGTATCATGTCAGAAAGCCTTAC 1090
```

RESULT 6

```
US-09-533-029-121
; Sequence 121, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER FILING DATE: 60/125,814
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 121
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G501
US-09-533-029-121
Alignment Scores:
Pred. No.: 1 42e-42 Length: 1216
Score: 462.00 Matches: 96
Percent Similarity: 59.8% Conservative: 38
Best Local Similarity: 42.9% Mismatches: 66
Query Match: 22.7% Indels: 24
Gaps: 3
```

```
US-10-780-703-2 (1-379) x US-09-533-029-121 (1-1216)
QY 39 AsnGluAspGlu-----AlaAspAspHisAspMet----- 50
Db 17 AACGAAGAAGATTTTTCAGAGCAACAGACAAGAAGAGATGAAGTCGAGCTA 76
QY 51 ---ValMetProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeu 69
Db 77 AATTACAGCTGGGTTCGGATTCCATCCAAACGAGCGAGGAGCTTGGAATCTACTTG 136
QY 70 ArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyr 89
Db 137 TGCAGAAATGTCTTCCGAGCAGATCTCGGCTCCGGTTCATCCGAGATTCCTCTAC 196
QY 90 ArgTyrAspProTrpGluLeuProAlaMetAlaAlaIleGlyLysGluTyrPhe 109
Db 197 AGTTCAATCCTTGGGAGCTTCAGAGATGTCCTCTACGAGAGAAAGATGGTACTTC 256
QY 110 TyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSer 129
Db 257 TTCTACCTAGAGATCGAAATACCCAAACGGTTCGGTCTTAACCGGCGAGGAAACC 316
QY 130 GlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIle 149
Db 317 GGTATTGAAAGCTACCGAGCAGATAAACCGATTGGTAAA-----CCGAAGACGTTG 370
QY 150 GlyLeuLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSer 169
Db 371 GGTATCAAGAAAGCACTCGTCTTCTACGAGGAAAGCTCCAAAGGGGATTAGACCAAT 430
QY 170 TrpIleMetAsnGluTyrArgLeuProHis-----HisGluThr 182
Db 431 TCGATTAATGATGAGTATCGTCTCGCTAATGTTGATAGATCAGCTTCTGTTAACAAG 490
QY 183 GluLysTyrGlnLysAlaGluLeuSerLeuCyArgValTyrLysArgProGlyValGlu 202
Db 491 AACAACTACGACTTCGATGATGGGTTTATGTGCGAATATACAAAGAAAGAACCATG 550
QY 203 Asp-----HisProSerValProArgSerLeuSerThrArgHisAsnHisAsnSer 220
Db 551 GAGAAATATTTCCCGCGGATGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 610
QY 221 Ser-----ThrSerSerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSer 236
Db 611 TCACCTTTTGATACATCAGACTTACCCGACATTGCAAGAGGATGATTCAGGAGC 670
QY 237 SerSerAsnHis 240
Db 671 TCAGGTGGTCAC 682
```

RESULT 7

```
US-09-533-029-23
; Sequence 23, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
```



```
Db 491 AAAACCGATTGGGTATCCGTCACGAATTCGTCCTCAAGGATCTCATCATCTCCCAATCAT 550
QY 181 GluThrGluLysTyrglnLysAlaGluLeuSerLeuCysArgValTyrlsArg----- 198
Db 551 TCCTGAGCTCT---CCAAAGGAAGACTGGGTCTTGTAGGTATTCATTAAGATAACG 607
QY 199 ProGlyVal-----GluAspHis-----ProSer 206
Db 608 GAAGGAGTTATATCTAGACACAACATGGGAAGCTGTTTTGATGAGACAGCCTCTGCGATCG 667
QY 207 ValProArgSerLeuSerThrArgHisAsnHisAsnSerSerThrSerSerArgLeu 226
Db 668 CTCCTCT---CCACTGATGATCTTACATCAACTTTGACCAAGAACCCCTCTTCTATCTC 724
QY 227 AlaLeuArgGlnGlnHisHisSerSerSerSerAsnHisSerAsp-----AsnAsn 244
Db 725 AGT-----GATGATCATCTACATCATCATGAGCAGCTACCCCTCTCTCCAAAT 775
QY 245 LeuAsnAsnAsnAsnAsnLeuLeuAsnAsnLeuGluLysLeuSerThrGluTyrlsArg 264
Db 776 TTCTCAGAGAACCACCACTTAAAC----- 799
QY 265 GlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 284
Db 800 -----TCGAACCTTAACCACTCACTCTCTGAACCTCAAGATTCCATGCAAGAAC 847
QY 285 GlnAsnIleTyrlsArgProMetProTyrlsArgProMetProTyrlsArgProMetPro 304
Db 848 CTTAAC-----CCCTTGTGTTTCTGGTGGTTCAGCCCTCAGCCAGCTCAGCAGGCTC 898
QY 305 ArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsn 324
Db 899 GACTCATTCTGTTCTTCAGATCAGATGTTCTCAGAGCTCTACTCAGTCACTCACTAAG 958
QY 325 TyrlsGlnIleSerAspGlyAlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMet 344
Db 959 ATTGATGGAAGCCCTCGGGCCCTAAGAATCATCAGAGTTATGGAAGAGTTCGAGAGC 1018
QY 345 AsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIle 364
Db 1019 CTCCTCAGCAGATCGGT---ATTCCAGCACT-----GTTTGAATG-CTG 1062
QY 365 ValProAspGlyAsnArgAspHisTyrlsArgHisTyrlsArgHisTyrlsArgHisTyrls 377
Db 1063 ATGATCGAGTGTAAACGAGATTACTATTGCTATATTCCT 1101

RESULT 9
US-09-533-029-73
; Sequence 73, Application US/09533029
; Patent No. 666446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; EARLIER FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
```

```
; LENGTH: 913
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G511
US-09-533-029-73

Alignment Scores:
Pred. No.: 8 45e-31 Length: 913
Score: 357.00 Matches: 75
Percent Similarity: 61.4% Conservative: 27
Best Local Similarity: 45.2% Mismatches: 46
Query Match: 17.6% Indels: 18
DB: 3 Gaps: 5

US-10-780-703-2 (1-379) x US-09-533-029-73 (1-913)
QY 49 AspMetValMetProGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyr 68
Db 37 GATGAGGTCACAATCGGTTTCGCTTCTATCCACGAAAGAACTGGTTTCGTTCTCTAC 96
QY 69 LeuArgArgLysValGluGlyLysArgPheAsnVal-----GluLeuIleThrPheLeu 86
Db 97 CTACGAAACAGCTCGAAGGAAGAGTGAATCAATGATCGTGTCTATCCCGTACTT 156
QY 87 AspLeuTyrlsArgTyrlsArgProTrpGluLeuProAlaMetAlaIle-----Gly 103
Db 157 GAGGCTTTGAGTTCGAGCTAGTCACTCTCCAAATGTTGCTGGAGTGAGATGTCGAGGA 216
QY 104 GluLysGlu---TrpTyrlsPheTyrlsValProArgAspArgLysTyrlsArgAsnGlyAspArg 122
Db 217 GACGCTGAGCAATGGTTCCTTCTTCGTGTCACGACAAACGCGCAAGCAGAGGAGGCGA 276
QY 123 ProAsnArgValThrThrSerGlyTyrlsTrpLysAlaThrGlyAlaAspArgMetIleArg 142
Db 277 CCGAGTAGAACTACTGGTTTCAGATACTGGAAGCAACTGGATCCTCGTGTCCAGTCTTT 336
QY 143 SerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrlsArgLysAla 162
Db 337 TCCAAAGACAACAATAATGATTGGAGCAAGAAACTATGTTTCTTACACTGGAAGCA 396
QY 163 ProLysGlyThrArgThrSerTrpIleMetAsnGluTyrlsArgLeuProHisHisGluThr 182
Db 397 CCCACAGGAAGAAACTAAATGGAATAATGAATGAGTAC-----CACGCCGTT 444
QY 183 GluLysTyrlsGlnLysAla-----GluIleSerLeuCysArg 194
Db 445 GACGAAACAGTCAACGCTTCCCAATCCCTAAGCTGAGAGCTGAGTTTATGTCGA 504
QY 195 ValTyrlsArgProGly 200
Db 505 GTCTACATAACAACAGGA 522

RESULT 10
US-09-533-029-115/c
; Sequence 115, Application US/09533029
; Patent No. 666446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
```



```

Db 926 AATCGGGCTGCAACATCGGTTATTGGAAGCTACCGAACCGACGCGTTTACT 985
Qy 143 SerCluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSerGlyLysAla 162
Db 986 TCCGGTGAACACAAAAGTTGGGGTAAAAAAGCGCTCGGTTTTTATACGGCGGTAAACCA 1045
Qy 163 ProlysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHisGluThr 182
Db 1046 CCAAAAGGGTAAAACTAATGGATCGTCATGATACAGAGTTGTAGAAAATTAACA 1105

```

RESULT 12

```

US-09-533-029-99
; Sequence 99, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Mareha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 99
; LENGTH: 1339
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G515
US-09-533-029-99

```

```

Alignment Scores:
Pred. No.: 3,94e-17 Length: 1339
Score: 239.50 Matches: 82
Percent Similarity: 45.7% Conservative: 50
Best Local Similarity: 28.4% Mismatches: 100
Query Match: 11.8% Indels: 58
DB: 3 Gaps: 11

```

US-10-780-703-2 (1-379) x US-09-533-029-99 (1-1339)

```

Qy 45 AspHisAspHisAspMetValMetPro---GlyPheArgPheHisProThrGluGluGlu 63
Db 139 GATCGCTGCAGAGAAATGGAACTCCCTGTGGTTTAAAGATTCTCGCCAGCCGACGAGGAG 198
Qy 64 LeuIleGluPheTyrLeu-ArgArgLys-----ValGluGlyLysArgPheAenValG1 81
Db 199 ATCGTCTCGATTAACCTTTGGCCGAAAAATTCGATAGACACGACGACCATGTCGATCGA 258
Qy 81 uLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAlaAl 101
Db 259 TTCATTAAACACA-GTCCCTGCTGTAGACTCGATCTTGGGAGTTACCTTGCCAGTCAAG 317
Qy 101 alleGlyLysGlu-----TrpTyrPheTyrValProArgAspArgLysTyrArgAs 119
Db 318 GATCAAACTGAAGATGTGGTCTGTTCTTCAGACCTTAAGGAGAACAAATATGGCAG 377
Qy 119 nGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspAr 139

```

```

Db 378 AGTGATCAGCAGATGAGAAAAACGAAATCTGGGTTTGGGAAGAGTACTGCGACACAAA 437
Qy 139 gMetIleArgSerGluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSe 159
Db 438 GCCTATC---ATCGTAATCGCAACAGATCGGTGAGAAAAAGATTGTGATGTTTACAC 494
Qy 159 rGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHi 179
Db 495 GACTAGGAATCCAAA-----TCCGATTGGTTATACACGAGTAC-----CA 536
Qy 179 sHisGluThrGluLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgPr 199
Db 537 CGGTTCTCTCATAAACCATGATGATGATACATACACTCTGTAAGTTATGTTTAATGG 596
Qy 199 oGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAs 219
Db 597 TGGCATG-----AGAGAGAAGTCTTCTCTC 620
Qy 219 nSerSerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerAs 239
Db 621 TTCTCTCTCTCTCTCTGGTGTAGTGAATTGAGCAGAGTCGTCGTGACTCTTTAATC-- 678
Qy 239 nHisSerAspAsnAsnLeuAsnAsnAsnLeuAsnLeuLysLeuLysLeuSerTh 259
Db 679 -----CCTCAGCTTGTCAACAATCTGAGGGATCCTCACTTCACAGAGAAGATCC 728
Qy 259 rGluTyrSerGlyAspGlySerThrThrThrThrThrAsnSerAsnSerAspValTh 279
Db 729 AAGTCAGTTTGGTGAT----- 744
Qy 279 rIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnTh 299
Db 745 -----GTGTCGCAAGAAGTCCCAATCGAGGATCGTAAACTGAC 782
Qy 299 r-----LeuIleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValas 317
Db 783 CGAGGAATGGTAAATGGCTGATGAATGATGAGGATGAT-----GCTCAAAATCGA 833
Qy 317 pAsp-LeuGlnArgLeuValasn 324
Db 834 GGATGCTATACCGATTGAGGAAT 856

```

RESULT 13

```

US-09-248-796A-336
; Sequence 336, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 336
; LENGTH: 1653
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-336

```

```

Alignment Scores:
Pred. No.: 0.000131 Length: 1653
Score: 131.00 Matches: 66
Percent Similarity: 33.8% Conservative: 41
Best Local Similarity: 20.8% Mismatches: 110
Query Match: 6.4% Indels: 100
DB: 3 Gaps: 12

```

US-10-780-703-2 (1-379) x US-09-248-796A-336 (1-1653)

Db 1291 GCTTCACGGATTTTA-----ACACACACATAACCCCAAGA 1329
 QY 338 rGlnAlaLeu 342
 Db 1330 TTTGGCGACAGTC 1342

RESULT 15
 US-09-270-767-12547
 ; Sequence 12547, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 12547
 ; LENGTH: 934
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 US-09-270-767-12547

Alignment Scores:
 Pred. No.: 0.000221 Length: 934
 Score: 125.50 Matches: 52
 Percent Similarity: 42.7% Conservative: 15
 Best Local Similarity: 33.1% Mismatches: 54
 Query Match: 6.2% Indels: 36
 DB: 3 Gaps: 7

US-10-780-703-2 (1-379) x US-09-270-767-12547 (1-934)

QY 219 AsnSerSerThrSerArgLeuAlaLeuArgGlnGlnHisSerSerSerSer 238
 Db 8 AACGGAGACGACGACG-----AGCAGCAGCAGCAGC 40
 QY 239 AsnHisSerAspAsnLeuAsnAsnAsnAsnAsnLeuGluLysLeuSer 258
 Db 41 AATAGTGGTAACACACACACACACACAAATACATCATCAAT----- 85
 QY 259 ThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThr 278
 Db 86 -----AACACTATACCAACACGACCAATAAATACCTGGTAACAAT 124
 QY 279 ThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsn 298
 Db 125 TCATCGGCC-----AAACCCCTACTATCGGCACCTACAGCTCGCTAAATCGG 172
 QY 299 ThrLeuIleValSerThrArgAsnHisGlnAspAspAsp---GluThrAlaIleValAsp 317
 Db 173 GTGCTC-----AATTCCGAGCGGATTCCCTGGAGATCCTGCAGCAGCAA 217
 QY 318 AspLeuGlnArgLeuValAsn-TyrGlnIleSerAspGlyAlaThrThrLeuMetProG1 337
 Db 218 CAACACAGCAGCAGCAGCTCAACGGGACCACTCAGCGCAACAGCACCACATCAATACCACA 277
 QY 337 nThrGlnAlaLeuAlaMetAsnMetIlePro-----AlaGlyThrIleProAsnAs 355
 Db 278 AGCGGGCGCAGCAGCTAGCTCGCGCAGATTCCACAGCAATCGGGACATAACTCACC 337
 QY 355 nAlaLeuTrpAspMetTrpAsnProIleValProAspGlyAsnArgAsp 371
 Db 338 GCCAATA-----GCAGTTCAACCAACGCGACCCGAGCAGGAAACGGGAA 380

Search completed: April 6, 2006, 10:02:43
 Job time : 249 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 09:54:37 ; Search time 4958 Seconds

(without alignments)
3576.503 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAIVSSTTSIIIPMSNQVNN.....MNPVIVDGNRHYTNIPK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q/abs/ABSWEB spool/US10780703/runat_06042006_101339_6361/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h
-USER=US10780703 @CGN_1_1_5315 @runat_06042006_101339_6361 -NCFU=6 -ICFU=3
-NO_MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1498	73.6	947	8 DR749909	DR749909 85-102025
C	1374	67.6	901	8 DR749910	DR749910 85-102025
3	1113	54.7	622	1 AV827617	AV827617 AV827617
4	882.5	43.4	880	8 DR938316	DR938316 EST112985
5	881.5	43.3	648	1 AW683672	AW683672 NF017D07L
6	878.5	43.2	653	2 BE204243	BE204243 EST396919
7	877	43.1	906	8 DR935181	DR935181 EST112672

8	871	42.8	853	8	DR928151
9	814.5	40.0	1404	10	CL960332
10	802.5	39.5	523	8	CX527114
11	799	39.3	516	3	B1893633
12	797.5	39.2	1245	10	CL971389
13	789.5	38.8	694	2	BF625246
14	782	38.4	525	8	DT006901
15	765	37.6	801	8	DR813028
16	761	37.4	784	6	CB621699
17	759.5	37.3	780	7	CN149464
18	756.5	37.2	763	6	CB649362
19	753.5	37.0	770	6	CB621762
20	740.5	36.4	555	1	AW443468
21	740.5	36.4	649	7	CK077495
22	737	36.2	742	8	DR787066
23	725.5	35.7	640	7	CN144198
24	712.5	35.0	657	2	B1268158
25	710	34.9	486	8	DT004645
26	707.5	34.8	741	7	CO096386
27	696.5	34.2	690	9	BH481683
C	688.5	32.9	813	9	AZ126200
28	668.5	32.7	494	7	CK076588
29	664.5	32.3	427	1	AV440745
30	657	32.3	468	1	AV442690
C	657	32.3	468	1	AV442690
C	644.5	31.7	823	9	CC659605
33	631.5	31.0	633	2	BG320780
C	631.5	31.0	879	9	CC689779
35	627	30.8	838	9	BH666975
36	618	30.4	879	9	CC694458
37	615	30.2	853	9	CC637855
C	614	30.2	929	9	CC637864
C	614	30.2	901	10	CG180706
C	612	30.1	973	9	CC714199
C	611	30.0	923	10	CG195779
42	605	29.7	859	10	CG180704
C	600.5	29.5	739	10	CL155695
44	594.5	29.2	540	8	DN183182
C	591	29.1	986	9	CC339647

ALIGNMENTS

DR749909	947 bp	mRNA	linear	EST 19-JUL-2005
85-L020254-065-002-E11-SeLA MP1Z-ADIS-065d				Arabidopsis thaliana
CDNA clone 002-E11, mRNA sequence.				
DR749909				
DR749909.1	GI:71035249			
EST.				
Arabidopsis thaliana (thale cress)				
Arabidopsis thaliana				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
1 (bases 1 to 947)				
Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B., Gilmarin, P., Giraudat, J., Parcy, F., Reindl, A., Sablowski, R., Coupland, G., Martin, C., Angenot, G.C., Baumeister, H., Mock, H.P., Carbonero, P., Colombo, L., Tonelli, C., Engstrom, P., Droege-Laser, W., Gatz, C., Kavanagh, T., Kushnir, S., Zabeau, M., Laux, T., Hordworth, M., Ruberti, I., Ratcliff, F., Smeekens, S., Somssich, I., Weisshaar, B. and Traas, J.				
REGIA, an EU project on functional genomics of transcription factors from Arabidopsis thaliana				
Comp. Funct. Genomics 3 (2), 102-108 (2002)				
Contact:				
Paz-Ares, Costantino, Vittorioso, Davies, Gilmarin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenot, Baumeister, Carbonero, Colombo, Tonelli, Engstrom, Droege-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Laux, Hordsworth, Ruberti, Smeekens, Somssich, Weisshaar, Traas				
Bielefeld University, Institute for Genome Research				
Universitaetsstrasse 25, D-33594 Bielefeld, Germany				

Email: bernd.weishaar@uni-bielefeld.de
 AGI: AFZG02450; SeqAnalysis: undetermined frame; Translation: no full cds detected
 Data analysis performed in the frame of REGULATORS (Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance), a Trilateral Co-Operation in Plant Genomics between Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent Thureau (IBP-Orsay UMR8618 CNRS-UPS, thureau-ad-ibp.u-psud.fr) and Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE, lecharny-ad-ibp.u-psud.fr).
 Definition of the terms used to describe the quality of the clone: The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGRv5 annotation, and the resulting AGI code is presented if more than 90 percent identity was found. The sequences were also blasted against all TIGRv5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for SeqAnalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGRv5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGRv5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.
 Insert Length: 947 Std Error: 0.00
 Seq primer: SeLA TCGCGTTACGCTAGCATGCTC.
 Location/Qualifiers
 1..947
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="GABI:1452250"
 /db_xref="taxon:3702"
 /clone="002-E11"
 /lab_host="E. coli DH5alpha"
 /clone_lib="MP1Z-ADIS-065d"
 /note="Vector: pDONR201; In the context of the EU-funded project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various A. thaliana tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG_tag):
 Spr-AATTCACGCTCACACC-3pr; RG_tag2:
 Spr-CATCGCAATCCCGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weishaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by Gabipd (http://gabi.rzpd.de)."

FEATURES
source

Pred. No.:	2,77e-146	Length:	947
Score:	1498.00	Matches:	277
Percent Similarity:	99.6%	Conservative:	3
Best Local Similarity:	98.6%	Mismatches:	1
Query Match:	73.6%	Indels:	0
DB:	8	Gaps:	0

US-10-780-703-2 (1-379) x DR749909 (1-947)

Qy	13	MetSerAsnGlnValAsnAsnAsnGluLysGlyIleGluAspAsnAspHisArgGlyGly	32
Db	104	ATGAGTAACCAAGTCAACAATAACGAAAAAGGTATAGACATATGATCATAGAGCGGC	163
Qy	33	GlnGluSerHisValGlnAsnGluAspGluAlaAspHisAspHisAspMetValMet	52
Db	164	CAGAGAGTATGTCCTCAAAATGAAGTGAAGTATGATCATGATCATGATGATGATG	223
Qy	53	ProGlyPheArgPheHisProThrGluGluGluLeuGluPheThrLeuArgArgLys	72
Db	224	CCCGATTAGATTCCATCTTACCGAAGAAGAACTCATAGAGTTTACCTTCGCGGAAA	283
Qy	73	ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuThrArgTyrAsp	92
Db	284	GTTGAGGCAACGCTTTAATGTAGAACTCATCTTCTCGATCTTTATCGCTATGAT	343
Qy	93	ProThrGluLeuProAlaMetAlaAlaIleGlyGluLysGluTyrPheTyrValPro	112
Db	344	CTTTGGGAATCTTCTGCTATGGCGCGATAGGAGAGAAAGTGTACTTCTATGTGCA	403
Qy	113	ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp	132
Db	404	AGAGATCGGAATATAGAAATGGAGATAGACCGAAGCAGTAACGACTTCAGGATATGG	463
Qy	133	LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys	152
Db	464	AAAGCCACCGAGCTGATAGATGATCAGATCGAGACTTCTCGGCTATCGGATTAAG	523
Qy	153	LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrIleMet	172
Db	524	AAACCCCTAGTTTCTACTCTGGTAAAGCCCTTAAAGGCACTCTGACTAGTTGGATCATG	583
Qy	173	AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeu	192
Db	584	AACGAGTATCGTCTTCGTACCAATGAACCGAGAAGTACCAAAAGGCTGAAATATCATTTG	643
Qy	193	CysArgValTyrLysArgProGlyValGluAspHisProSerValProArgSerLeuSer	212
Db	644	TGCGAGTGTACAAAGCCAGGAGTAAAGATCATCATCGGTACCACTCTCTCTCTCC	703
Qy	213	ThrArgHisAsnHisAsnSerThrSerArgLeuAlaLeuArgGlnGlnGln	232
Db	704	ACAGATCATATAACCACTAATCTCATCGACATCATCCGTTTAGCTTAAGACACACAA	763
Qy	233	HisHisSerSerSerSerAsnHisSerAspAsnLeuAsnAsnAsnAsnAsnIleAsn	252
Db	764	CACCATTCCTCTCTCTAATCATTTCCGACCAACCAACCTTAACCAACCAACCAATTCAC	823
Qy	253	AsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThr	272
Db	824	AATCTCAGAGAGCTCTCCACCGAATATTCGCGCGACCGGACGACAAACACACACCA	883
Qy	273	AsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetPro	292
Db	884	AACAGTAACCTCTGACGTTACCATTTGCTTAGCCCATCAAAACATATATATCGTCCATGCT	943
Qy	293	Tyr	293
Db	944	TAC	946

RESULT 2

DR749910/c

LOCUS

DEFINITION

85-L020255-065-002-E11-SeLB MP1Z-ADIS-065d Arabidopsis thaliana

ORIGIN

Alignment Scores:

DR749910 901 bp mRNA linear EST 19-JUL-2005
 LOCUS 85-L020255-065-002-E11-SeLB MP1Z-ADIS-065d Arabidopsis thaliana

```

cdna clone 002-E11, mRNA sequence.
DR749910
DR749910.1 GI:71035250
EST
SOURCE
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 901)
Paz-Ares,J., Valencia,A., Costantino,P., Vittorioso,P., Davies,B.,
Gilmartin,P., Giraudat,J., Parcy,F., Reindl,A., Sablowski,R.,
Coupland,G., Martin,C., Angenent,G.C., Baeumlein,H., Mock,H.P.,
Carbonero,P., Colombo,L., Tonelli,C., Engstrom,P.,
Droeghe-Laser,W., Gatz,C., Kavanagh,T., Kuehnir,S., Zabeau,M.,
Laux,T., Hordeworth,M., Ruberti,I., Ratcliff,F., Smeekens,S.,
Somssich,I., Weissshaar,B. and Traas,J.
REGIA, an EU project on functional genomics of transcription
factors from Arabidopsis thaliana
Comp. Funct. Genomics 3 (2), 102-108 (2002)
Contact:
Paz-Ares, Costantino, Vittorioso, Davies, Gilmartin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenent, Baeumlein, Carbonero, Colombo, Tonelli, Engstrom, Droeghe-Laser, Gatz, Kavanagh, Kuehnir, Zabeau, Laux, Hordeworth, Ruberti, Smeekens, Somssich, Weissshaar, Traas
Bielefeld University, Institute for Genome Research
Universitaetsstrasse 25, D-33594 Bielefeld, Germany
Email: bernd.weissshaar@uni-bielefeld.de
AGI: AT2002450; SeqAnalysis: undetermined frame; Translation: no full cds detected
Data analysis performed in the frame of REGULATORS (Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance), a Trilateral Co-Operation in Plant Genomics between Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent Thareau (IBP-Orsay UMR8618 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and Alain Lecharny (URGV-EVry UMR INRA-CNRS-UEVE, lecharny-ad-ibp.u-psud.fr).
Definition of the terms used to describe the quality of the clones:
The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'Contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGRv5 annotation, and the resulting AGI code is presented if more than 90 percent identity was found. The sequences were also blasted against all TIGRv5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for SeqAnalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGRv5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGRv5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.
Insert Length: 901 Std Error: 0.00
Seq primer: Seub GTAACATCAGAGATTGTGAGAC.
Location/Qualifiers
1. 901
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:1452251"
/db_xref="taxon:3702"
/clones="002-E11"
/lab_host="E. coli DH5alpha"

```

/clone lib="MPIZ-ADIS-065d"
/notes=Vector: pDONR201; In the context of the EU-funded project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various *A. thaliana* tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG_tag1: 5pr-AATTCAGCTCCACC-3pr; RG_tag2: 5pr-CATGCAATTCGGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weissshaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by GabiPD (<http://gabi.rzpd.de>).

ORIGIN

Alignment Scores:	2.66e-133	Length:	901
Pred. No.:	1374.00	Matches:	259
Score:	98.9%	Conservative:	2
Best Local Similarity:	98.1%	Mismatches:	3
Query Match:	67.6%	Indels:	0
DB:	8	Gaps:	0

US-10-780-703-2 (1-379) x DR749910 (1-901)

QY	113	ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp	132
DB	901	AGAGATCGGAATATAGAAATGAGATAGACCGACCGAGTACGACTTCAGGATATTGG	842
QY	133	LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys	152
DB	841	AAAGCCACCGAGCTGATAGGATGATCAGATCGGAGACTTCTCGGCCCTATCGGATTAAAG	782
QY	153	LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet	172
DB	781	AAAAACCTTAGTTTCTACTCTGTGTAAGCCCCCTAAAGGCACCTCGTACTAGTTGGATCATG	722
QY	173	AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeu	192
DB	721	AACGAGTATCGTTCCTCGTACCATGAACCGAGAGTACCACCAAGGCTGAAATATCATTTG	662
QY	193	CysArgValTyrLysArgProGlyValGluAspHisProSerValProArgSerLeuSer	212
DB	661	TGCCGAGTGTACAAAAGGCCAGGAGTAGAAGATCATCCATCGGTACCCAGCTTCTCTCTCC	602
QY	213	ThrArgHisHisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnGln	232
DB	601	ACAAGACATCATCAACATACTCATCGACATCATCCGTTTAGCTTTAAGACACACAA	542
QY	233	HisHisSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	252
DB	541	CACCATTCATCT	482
QY	253	AsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThr	272
DB	481	AATCTCGAGAAGCTCTCCACCGAATATTCCGGCGAGCGGACGACCAACCAACGACCA	422
QY	273	AsnSerAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	292
DB	421	AACAGTAACTTGACGCTTACCATTTGCTTAGCCCAATCAAAACATATATATCGTCAATCCCT	362
QY	293	TyrAspThrSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	312

FEATURES
source

```

|||||
Db 361 TACGACACAGCAACACACATGTCCTACGAGAAATCATCAAGACGATGAA 302
QY 313 ThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGlyAlaThr 332
Db 301 ACTGCCATTGTTGACGATCTTCAAGAGACTAGTTAACTACCAATATCAGATCGACGACA 242
QY 333 ThrLeuMetProGlnThrGlnAlaLeuAlaMetAsnMetIleProAlaGlyThrIle 352
Db 241 AGCCTAATGCTCAAACTCAAGCGGCTTAGCTATGAACATGATTCCTGCGAGAACGAAT 182
QY 353 ProAsnAsnAlaLeuTrpAspMetTrpAsnProIleValProAspGlyAsnArgAspHis 972
Db 181 CCAACAATGCTTTGGGGGATATGTGGGATCCCAATAGTACCAGATGGAAACAGAGATCAC 122
QY 373 TyrThrAsnIle 376
Db 121 TATACATAACATG 110

RESULT 3
AV827617 622 bp mRNA linear EST 01-APR-2002
LOCUS AV827617 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-17-121 5',
DEFINITION mRNA sequence.
ACCESSION AV827617.1 GI:19869677
VERSION AV827617
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 622)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda PLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and Sali. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source Location/Qualifiers
1. .622
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-17-121"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="PH108"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: Sali; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Alignment Scores: 3.89e-106 Length: 622
Pred. No.: 1113.00 Matches: 206
Score: 100.0% Conservative: 0
Percent Similarity: 100.0%
Best Local Similarity: 100.0% Mismatches: 0

```

```

Query Match: 54.7% Indels: 0
DB: 1 Gaps: 0

US-10-780-703-2 (1-379) x AV827617 (1-622)

QY 88 LeuTyrArgTyrAspProTrpGluLeuProAlaMetAlaIleGlyGluLysGluTrp 107
Db 3 CTTTATCGCTATGATCCTTGGAACTTCTCTGCTATGGCGCGATAGAGAGAGAGTGG 62
QY 108 TyrPheTyrValProAspAspArgLysTyrArgAsnGlyAspArgProAsnArgValThr 127
Db 63 TACTTTCATGTGCGCAAGAGATCGGAATATAGAATGGATAGACCGAACCGAGTAACG 122
QY 128 ThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArg 147
Db 123 ACTTCAGGATATTCGAAAGCCACCGAGCTCATAGGATGATCAGATCGAGACTTCTCGG 182
QY 148 ProIleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArg 167
Db 183 CCTATCGGATTAAGAAACCCCTAGTTTCTCTGTTAAAGCCCTTAAAGGCACTCGT 242
QY 168 ThrSerTrpIleMetAsnGluTyrArgLeuProHisGluThrGluLysTyrGlnLys 187
Db 243 ACTAGTTGGATCATGAACGAGTATCGTCTTCGACCATGAACCGAGAGTACCAAG 302
QY 188 AlaGluIleSerLeuCysArgValTyrLysArgProGlyValGluAspHisProSerVal 207
Db 303 GCTGAATATCATTTGTCGAGTGTACAAAGGCGAGGTAGAGATCATCCATCGGTA 362
QY 208 ProArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAla 227
Db 363 CCACGTTCTCTCTCCACAAGACATCAACCACTCATCGACATCATCCCGTTTACGC 422
QY 228 LeuArgGlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 423 TTAAGACAAACAAACCAACCATTCCTCTCTCTATATTCGACACACACCTTAAACAC 482
QY 248 AsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThr 267
Db 483 AACAAACACATCAACATCTCGAGAGCTCTCCACCGAATATTCGCGGCGAGGCACA 542
QY 268 ThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIle 287
Db 543 ACAACACAGCACACAAACAGTAACTCTGACGTTACCATTTGCTTAGCCCAATCAAAACATA 602
QY 288 TyrArgProMetProTyr 293
Db 603 TATCGTCCAATGCGCTTAC 620

RESULT 4
DR938316 880 bp mRNA linear EST 02-AUG-2005
LOCUS DR938316 Aquilegia cDNA library Aquilegia formosa x Aquilegia
DEFINITION pubescens cDNA clone COLPM65, mRNA sequence.
ACCESSION DR938316.1 GI:71707679
VERSION DR938316
KEYWORDS EST.
SOURCE Aquilegia formosa x Aquilegia pubescens
ORGANISM Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Aquilegia.
1 (bases 1 to 880)
Hodges,S.A., Rensink,W., Buell,C.R., Borevitz,J., Kramer,E.,
Nordborg,M. and Tomkins,J.
Generation of ESTs from Aquilegia
Unpublished (2005)
Other ESTs: EST1129854
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813

```

FEATURES
 source
 1. .880
 /organism="Aquila formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="CO1PM65"
 /tissue_type="mixed shoot and floral apical meristems, flower buds, leaves and roots"
 /lab_hosts="DH10B T1 (T1 and T5 phage resistance)"
 /clone_lib="Aquila cDNA library"
 /notes="vector: pCMV SPOR6.1; Site 1: EcoRI; Site 2: NotI; F2, F3, and F4 lines of Aquilegia formosa X A. pubescens were grown from seed in greenhouses at UC Santa Barbara. Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively), 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

Alignment Scores:
 Pred. No.: 9.92e-82 Length: 880
 Score: 882.50 Matches: 186
 Percent Similarity: 65.3% Conservative: 29
 Best Local Similarity: 56.5% Mismatches: 67
 Query Match: 43.4% Indels: 48
 DB: 8 Gaps: 7

US-10-780-703-2 (1-379) x DR938316 (1-880)

QY 38 GlnAsnGluAspGluAlaAspAspHisAspMetValMetProGlyPheArgPhe 57
 |||||
 DB 7 CAAAACACGACGACGAGGATACACGACGACGATGATGCTGCTGCTCCGTTT 66
 QY 58 HisProThrGluGluGluLeuLeuGluPheTyrLeuArgLysValGluGlyLysArg 77
 |||||
 DB 67 CACCCCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 126
 QY 78 PheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAspProThrGluLeuPro 97
 |||||
 DB 127 TTTAATGTGGAGCTTATTACTTCTCGATCTATATCGCTATGACCTCGGAGCTTCCA 186
 QY 98 AlaMetAlaAlaLeuGlyGlyLysGluTyrPheTyrValProArgAspArgLysTyr 117
 |||||
 DB 187 GCATTAGCTGCCATTGAGAGAAGAAGTGTCTTCTATGTGCACGAGACCGAAGATAT 246
 QY 118 ArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyVala 137
 |||||
 DB 247 CGCAACGGAGATCGCCCTAACCGGCTGACACATCTGGCTACTGGAAGCTACTGGTGA 306
 QY 138 AspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPhe 157
 |||||
 DB 307 GACAGAATGATCCGAAATGATACTTCAGATCCATTGGGCTTAAGAAGACACTAGTCTTC 366
 QY 158 TyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu 177
 |||||
 DB 367 TATTCGGGGAAGACACCTAAGGATCAGAACGATGGATGATGATGATGATGATGATGAT 426
 QY 178 ProHisGluThrGluLysTyrGlnLysAlaGluLysSerLeuCysArgValTyrLys 197
 |||||

Db 427 CCTCAGCAAGACACTGATAGATACCAAAAGGCGAGAAATTCGCTTGTGAGTCTACAG 486
 QY 198 ArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsn 217
 |||||
 Db 487 AGACCTGGAGTAGAAGACACGACGAGTCCCGGCTCC-----TCT 528
 QY 218 HisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSer 237
 |||||
 Db 529 CATCTCTTCACCATCT-----TCATCAAGA 555
 QY 238 SerAsnHisSerAspAsnAsnLeuAsnAsn-----AsnAsnAsnLeuAsnLeuGlu 255
 |||||
 Db 556 GGAACCTCAGCCGGATAAAAGACAGCAGAACTGATCATCTATGGAATTTCAAGTTTCAA 615
 QY 256 LysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThr 275
 |||||
 Db 616 GCATAGATAGGACAACTCACCAGCACTACTACTACTACTACTACTACTACTACTAGATG 675
 QY 276 SerAspValThrIleAlaLeuAlaAsn-----GlnAsnIleTyrArgProMetProTyr 293
 |||||
 Db 676 GAAAGCTGATGAAGCAGATGCAAGCAGTACTAGCAATGA-TACAGGAACCTGCTCTGG 734
 QY 294 AspThr-----SerAsnAsnThrLeu-----IleValSerThrArgAsnHisGln 308
 |||||
 Db 735 GATAGCAATGTCTTCTCAATGATCATGATTCGCGAGTATGTCACAGCTTCAACCATTTGAG 794
 QY 309 AspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSer 328
 |||||
 Db 795 GAGGTCAGTGGTACTGCTATTCATTCAACAATATTCAAGG----- 833
 QY 329 AspGlyAlaThrThrLeuMetProGlnThrGlnAlaLeuAlaLeuAlaMetAsnIlePro 348
 |||||
 Db 834 -----CAAATGAGTTCGCTACCA 851
 QY 349 AlaGlyThrIleProAsnAsnAlaLeu 357
 |||||
 Db 852 TCAAATACTTGTAAATACAAATACCCCTC 878
 RESULT 5
 AW683672
 LOCUS
 DEFINITION
 AW683672 648 bp mRNA linear EST 15-JUN-2000
 NF017D07LF1061 Developing leaf Medicago truncatula cDNA clone
 NF017D07LF 5', mRNA sequence.
 AW683672
 VERSION
 AW683672.1 GI:7558407
 KEYWORDS
 SOURCE
 ORGANISM
 Medicago truncatula (barrel medic)
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 648)
 TORRES-JEREZ, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
 Expresed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula leaf library
 Unpublished (2000)
 CONTACT: May GD
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 224 6650
 Fax: 580 224 6692
 Email: gdmay@noble.org
 Insert Length: 648 Std Error: 0.00
 Plate: 017 row: D column: 07
 Seq primer: TCACACGAGAAACAGCATGAC.
 Location/Qualifiers
 1..648
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"


```

Db      255 TATCGTATGACCTTGGAACTTCTGCTTTGGCAGCAATCGGAGAGAGAGAGTGTGAC 314
Qy      109 PheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThr 128
Db      315 TTTTATGTCCTCGAGATAGGAAGTATCGAACCGGTGATCGTCCAAATCGTGTAACT 374
Qy      129 SerGlyTyrTyrLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgPro 148
Db      375 TCTGGTATTGGAGGCAACAGGCGGTAGGATTCGAATCGAACTGAAACTTCGGTCA 434
Qy      149 IleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThr 168
Db      435 ATTGACTCAAGAAACCTAGTTTCTTCTTCTGGAAGCTCTTAAAGGCATCGAAC 494
Qy      169 SerTrpIleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAla 188
Db      495 AGTTGGATTATGAATGAGTATAGATTATCCCAACATGAACCGAATGATCAAAAGGGT 554
Qy      189 GluIleSerLeuCyAsnValTyrLysArgProGlyValGluAspHisProSerValPro 208
Db      555 GAGATATCCCTATGTCGTGTATACAGAGAGCAGGATAGAGATCATCTCCCTCCCT 614
Qy      209 ArgSerLeuSer---ThrArg-HisHisAsnHisAsn 219
Db      615 CGTGTCTCCCAATTACACGACCATCATCATCAT 651

RESULT 7
DR935181
LOCUS      DR935181
DEFINITION EST1126720 Aquilegia cDNA library Aquilegia formosa x Aquilegia
pubescens cDNA clone COLP234, mRNA sequence.
ACCESSION DR935181
VERSION    DR935181.1 GI:71704544
KEYWORDS   EST.
SOURCE     Aquilegia formosa x Aquilegia pubescens
ORGANISM   Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Aquilegia.
REFERENCE  1 (bases 1 to 906)
            Hodges,S.A., Rensink,W., Buell,C.R., Borevitz,J., Kramer,E.,
            Nordborg,M. and Tomkins,J.
            Generation of ESTs from Aquilegia
            Unpublished (2005)
            Other ESTs: EST1126719
            Contact: Scott Hodges
            Department of Ecology, Evolution and Marine Biology
            University of California, Santa Barbara
            Santa Barbara, CA 93106, USA
            Tel: 805 893 7813
            Fax: 805 893 4724
            Email: hodges@lifesci.ucsb.edu
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source            1..906
                     /organism="Aquilegia formosa x Aquilegia pubescens"
                     /mol_type="mRNA"
                     /db_xref="taxon:338618"
                     /clone="COLP234"
                     /tissue_type="mixed shoot and floral apical meristems,
                     flower buds, leaves and roots"
                     /lab_hosts="DH10B T1 (T1 and T5 phage resistance)"
                     /clone_lib="Aquilegia cDNA library"
                     /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
                     F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
                     were grown from seed in greenhouses at UC Santa Barbara.
                     From these plants three sets of tissue were collected: 1)
                     Small flower buds (<10 mm) and very young inflorescences
                     (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
                     large (at or near anthesis) flower buds (65 & 35% by
                     weight respectively) and 3) Shoot apical meristems. A
                     fourth set of tissue was collected from plants of A.

```

formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

```

Alignment Scores:      3.91e-81      Length:      906
Pred. No.:            877.00      Matches:      174
Score:                74.3%      Conservative: 25
Percent Similarity:   64.9%      Mismatches: 35
Best Local Similarity: 43.1%      Indels:      34
Query Match:         8           Gaps:        5
DB:

US-10-780-703-2 (1-379) x DR935181 (1-906)

Qy      38  GlnAsnGluAspGluAlaAspAspHisAspHisMetValMetProGlyPheArgPhe 57
Db      143 CAAAACAACGACGACGACGAGGATAACCCAGCACGATATGTCATGCTCGCTTCGGTTT 202
Qy      58  HisProThrGluGluGluLeuGluPheTyrIleuArgGlyValGluGlyLysArg 77
Db      203 CACCCAAACCAGGAGAAACATTCATCTCCGCCGTAAAGGTTGAGGGCAAGCGC 262
Qy      78  PheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTrpGluLeuPro 97
Db      263 TTTAATGTGGAGCTTATTACTTTCTCGATATATATCGCTATGACCTTGGAGCTTCCA 322
Qy      98  AlaMetAlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyr 117
Db      323 GCATTAGCTGCCATTGGAGAGAAAGATGGTCTTCTATGTGCCACGACGACGAAATAT 382
Qy      118 ArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAla 137
Db      383 CGTAACGAGATGCCCTAACCGGTGCAACATCTGGCTACTGAGAGGCTACTGGTGCA 442
Qy      138 AspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPhe 157
Db      443 GACAGATGATCCGAATGATACTTGAGATCCATTGGCTTAAAGAGACACTAGTCTTC 502
Qy      158 TyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu 177
Db      503 TATTCCGGAAAGCACCTAAAGGCATCAGAACCAAGTTGGATCATGAATGAGTACCGTTTA 562
Qy      178 ProHisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeuCyAsnValTyrLys 197
Db      563 CCTCAGCAAGACACTGATAGTACCAAAAGCGAAATTTCCGCTTGTTCGAGTCTACAG 622
Qy      198 ArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisAsn 217
Db      623 AGACCTGGATGAGAGACCCAGCCAGAGTTCCCGGCTCC-----TCT 664
Qy      218 HisAsnSerSerThrSerArgLeuAlaLeu-----ArgGlnGlnHisHis 234
Db      665 CATTCCTCGACCATCTTCATCAAGAGGAAGTCAAGCCGCGGATTAAGACAGACGACGAC 724
Qy      235 SerSerSerSerAsnHisSerAspAsnLeuAsnAsnAsnAsnAsnAsnAsn 254
Db      725 ATCATTATGGAATTCGAGCTTTCAAGCATACATAGTAGCAACCACTCA----- 772
Qy      255 GluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThr 272
Db      773 -----CCCGCAACTACTACTACTACTACTACTACTACTACTACTAGATG 805
Qy      273 -----AsnSerAsnSerAspValThrIleAlaLeuAla 283
Db      806 GAAAAGCTGAATGACGAGATGCAAGCAGTACTAGCAATGATATACCGGAACCTGCTCTTGGG 865

```


Db 49 GAGGATAATGATGATACGAGCAGCAGATGTCATGCCCGGCTTTCGCTTCCACCCAACT 108
 Qy 61 GluGluLeuLeuGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
 Db 109 GAAGAAGAGCTCGTGGATCTTACTCTCGCGTAAAGTGGAGGAAAGCGTTTCAACGTT 168
 Qy 81 GluLeuLeuThrPheLeuAppLeuTyrArgTyrAspProThrGluLeuProAlaMetAla 100
 Db 169 GAGCTTATTACTTTCCTCGATCTTATCGCTATGACCCCTGGGAGCTTCCCGCTTGCCA 228
 Qy 101 AlalLeuGlyLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGly 120
 Db 229 GCTATTGGTGAGAGAGTGGTACTTCTATGTGCCAGAGATAGAAAGTATGCAACGGT 288
 Qy 121 AspArgProAsnArgValThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet 140
 Db 289 GATGCTCCCATCGTGTACTACTCTGGGTATTTGGAAAGCAACAGAGCTGATAGATG 348
 Qy 141 IleArgSerGluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSerGly 160
 Db 349 ATCGGAACCGAGATTTTTCGCTCTATCGGCTCAAGAAACCCCTAGTTTCTATTCTGGG 408
 Qy 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis 180
 Db 409 AAAGCTCTTAAGGCATACGAACCAAGTGGATTATGAACGAGTATCGCTTGGCCCAACAT 468
 Qy 181 GluThrGluLysTyrGlnLysAlaGluLeuSerLeuCysArgValTyr 196
 Db 469 GAAACTGAACGATATCAAAAGGCTGAGATATCGCTTTCGCGGTGTATC 516

RESULT 12

CL971389

LOCUS

DEFINITION

OaIFCC021242 Oryza sativa Express Library Oryza sativa (indica

cultivar-group) genomic, genomic survey sequence.

CL971389

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1. 1245

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Express Library"

/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

1.36e-72

797.50

55.8%

44.7%

Length:

Matches:

Conservative:

Mismatches:

1245

185

46

88

Query Match: 39.2%

DB: 10

Indels: 95

Gaps: 16

US-10-780-703-2 (1-379) x CL971389 (1-1245)

Qy 32 GlyGlnGluSerHisValGlnAsnGluAspGluAla-----AspAspHis 46

Db 40 GGGAGGAAGAGCGGCGCGCATAGGAGGAGCGCGCGCGCTGTGCGCGCGCATCGCAC 99

Qy 47 AspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGluLeuGlu 66

Db 100 GAGAACGACTGTGTATGCGGGTTCGGTTCCACCCGAGCGAGGAGGAGTGTATCGAG 159

Qy 67 PheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeu 86

Db 160 TTCTACTCCGGCGGAAGGTGGAGGCGCGCGGTTCACCTCGAGCTCATCACCCTCTC 219

Qy 87 AspLeuTyrArgTyrAspProThrGluLeuProAlaMetAlaAlaLeuGlyGluLysGlu 106

Db 220 GACCTCTACCGCTTCGACCCATGGAGCTCCCGCAATGGCGGTGATTGGGGAAGAG 279

Qy 107 TrpTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgVal 126

Db 280 TGGTTCTTTTACGTGCGCGGACCGCAAGTACCGAACCGCGCCGACCGCGGTG 339

Qy 127 ThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSer 146

Db 340 ACGCGCTCGGGTACTTGAAGCGCGGGCGCGCGATGATCGCGCGGAGAACAGC 399

Qy 147 ArgProIleGlyLeuLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThr 166

Db 400 CGCCCATCGGCTCAAGAAGACGCTCGTCTTCTACTCCGCGAAGGCCCAAGGGCGTC 459

Qy 167 ArgThrSerTrpIleMetAsnGluTyrArgLeu-----ProHisHisGluThrGluLys 184

Db 460 CGCAGCAGCTGGATCATGAACGATACCGCTTCCCGCGCGCGAGCGCGATCTC 519

Qy 185 TyrGlnLysAlaGluLeuSerLeuCysArgValTyrLysArgProGlyValGluAsp--- 203

Db 520 TTCTACAGTCTGAGATCTCGCTCTCGCGGTCTACAAGCGCTCCGCGCATCGACGCGC 579

Qy 204 -----HisProSerValPro-----ArgSerLeuSerThr 213

Db 580 CACGGCCACCCACGCGCGCGCGCGCAACGTCGTCGTCGTCGCGCGCGCGCGCG 639

Qy 214 ArgHisHisAsnHisAsnSerThrSerArgLeuAlaLeuArgGlnGlnHis 233

Db 640 AGCGCGCGGAGGAGCAGCTCCGGCAACACACTGCTGCCGGCTTACCGGCTTCCGGCAT 699

Qy 234 HisSerSerSerSerHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsn 253

Db 700 AGCGCGTCCCGTCTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 759

Qy 254 LeuGluLysLeuSerThrGluTyrSerGlyAspGlySer----- 266

Db 760 CTG-----CTCAAGGCGAGTGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 813

Qy 267 -----ThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 284

Db 814 CTCCAGCATCGGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 852

Qy 285 GlnAsnIleTyrArgProMetPro----- 292

Db 853 CAGCTGTGTATGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 912

Qy 293 -----TyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGln----- 308

Db 913 GCGGCG 972

Qy 309 -----AspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuVal----- 323

Db 973 GCG 1032

Qy 324 -----AsnTyrGlnIleSerAspGlyAla----- 331
 Db 1033 AGCCATGCGCGCTACGGTAATACACATTTGCGCGGAAGCCATCATCTCC 1092
 Qy 332 -----ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIlePro 348
 Db 1093 TTGCGCGCTGCACCTGATGCGGAG----- 1119
 Qy 349 AlaGlyThrIlePro-----AsnAsnAlaLeuTyrAspMet 360
 Db 1120 CTGGCAGCTGCCGATCTCGCGCGCTGGCGCGCTCGACAGGCTCTGGGAT--- 1176
 Qy 361 TrpAsnProIleValProAspGlyAsn-----ArgAspHis 372
 Db 1177 TGGAGCTCG---GTCCTGCACACGTCGACGGCCAGGATTAC 1215

RESULT 13
 BF625246 694 bp mRNA linear EST 18-OCT-2001
 LOCUS HVSM0008A15f Hordeum vulgare seedling shoot EST library
 DEFINITION HVSM0008A15f (Cold stress) Hordeum vulgare subsp. vulgare cDNA clone
 HVSM0008A15f, mRNA sequence.
 ACCESSION BF625246
 VERSION BF625246.3 GI:16256388
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 694)
 AUTHORS Wing, R., Close, T.J., Kleinohs, A., Wise, R., Begum, D., Frisch, D.,
 Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R.,
 Choi, D.W., Fenton, R.D. and Main, D.
 TITLE Development of a genetically and physically anchored EST resource
 for barley genomics: Morex cold-stressed seedling shoot cDNA
 library
 JOURNAL Unpublished (2001)
 COMMENT On Dec 18, 2000 this sequence version replaced gi:13082400.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864.656.7288
 Fax: 864.656.4293
 Email: rwing@clemson.edu
 Total hq bases = 583
 Seq primer: AATTAACTCTACTAAGGG
 High quality sequence stop: 691.
 Location/Qualifiers
 1..694
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Morex"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="HVSM0008A15f"
 /tissue_type="Seedling shoot"
 /lab_host="TJCI21"
 /clone_lib="Hordeum vulgare seedling shoot EST library
 HVSM0001 (Cold stress)"
 /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
 Seeds were surface sterilized then germinated under axenic
 conditions in the dark at room temperature on filter paper
 with water, nystatin and cefotaxime in covered
 crystallization dishes. Five-day old seedlings were
 incubated at 50C for 2 days. Shoots were then harvested,
 total RNA was prepared, poly(A) RNA was purified, one
 primary unamplified cDNA library was made, and 60000 pfu
 were in vivo excised to give pBluescript SK(-) cDNA
 phagemids. These steps were performed in the TJ Close
 laboratory at the University of California, Riverside
 (Choi, Close, Fenton). Phagemids were plated and picked at

the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu> see Close TJ, Wing R, Kleinohs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

ORIGIN

Alignment Scores:
 Pred. No.: 4,08e-72 Length: 694
 Score: 789.50 Matches: 144
 Percent Similarity: 91.8% Conservative: 12
 Best Local Similarity: 84.7% Mismatches: 11
 Query Match: 38.8% Indels: 3
 DB: 2 Gaps: 1

US-10-780-703-2 (1-379) x BF625246 (1-694)

Qy 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGlu 63
 Db 159 GACGCGCAGGAGGAGGACCTGGTGTATCCCGGGGTTCCTCCACCGAGGAGGAG 218
 Qy 64 LeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeu 83
 Db 219 CTCATCAGGTTTACCTCCCGCCGCAAGTGGAGGCGAGCGCTTCATGTGAGCTCATC 278
 Qy 84 ThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAlaAlaGly 103
 Db 279 ACCTTCCTCGACCTCTACCGCTACGACCCCTGGGAGCTCCCGAGCGTGGCAGCATTCGG 338
 Qy 104 GluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgPro 123
 Db 339 GAGAAGAGGTGGTTCTTCTACGTGCCGAGGACCGCAAGTACCGGAGACGGGACCGGCC 398
 Qy 124 AsnArgValThrThrSerGlyTyrTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSer 143
 Db 399 AACCGGTGACGGCTGGGGTACTGGAGCGGACGGGGCGGACAGATGATCCGGGCG 458
 Qy 144 GluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSerGlyLysAlaPro 163
 Db 459 GAGAGCAGCGCCCATCGGCTCAAGAAGACGCTCGTCTTCTACTCCGCAAGCGGCC 518
 Qy 164 LysGlyThrArgThrSerTyrIleMetAsnGluTyrArgLeuProHisHisGluThrGlu 183
 Db 519 AAGGCGTCCGCGACAGCTGGATCATGAACGAGTACCGGCTCCCGACCGCGACACCGAC 578
 Qy 184 LysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGlyValGluAsp 203
 Db 579 CGATACACAGACAGACAGAACTCTCTCTGCCCGCTTACAGCGCAGCGGATCGACGAC 638
 Qy 204 -----HisProSerValProArgSer 210
 Db 639 GGCGCGCGCCACCCCTCTCTNCGGGCGGTCTG 668

RESULT 14
 DT006901
 LOCUS DT006901
 DEFINITION DT006901 525 bp mRNA linear EST 05-AUG-2005
 clone VVG039B06 758717 Cabsau Cell Culture (CEL00001) Vitis vinifera cDNA
 DT006901
 ACCESSION DT006901.1 GI:71857846
 VERSION DT006901.1
 KEYWORDS Vitis vinifera
 SOURCE Vitis vinifera
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Score: 765.00 Matches: 155
Percent Similarity: 66.7% Conservative: 29
Best Local Similarity: 56.2% Mismatches: 57
Query Match: 37.6% Indels: 36
DB: 8 Gaps: 5

US-10-780-703-2 (1-379) x DR813028 (1-801)

QY 23 GlyIleGluAspAsnAspHisArgGlyGlnGluSerHisValGlnAsnGluAspGlu 42
DB 32 GCGACACCGCAGCAGCAGCGCGCGCGGACGATGGGCACCTGCAGCAAGCGCGCGGA 91
QY 43 AlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGlu 62
DB 92 GGC-----GACATGGTGATGCCGGGGTTCCGGTTCACCCCGGAGGAG 136
QY 63 GluLeuIleGluPheTyrLeuArgArgValGluGlyLysArgPheAsnValGluLeu 82
DB 137 GAGCTCATCGACTTCTACCTCCGCCGTAGGTGGAGGCAAGCGTTTCAACATCGAGCTC 196
QY 83 IleThrPheLeuAspLeuTyrArgTyrAspProTTPGluLeuProAlaMetAlaIle 102
DB 197 ATCAACCTCGTGGACCTCTACCGCTACGACCCATGGGATCTCCCTGCTTGGCTTCGATC 256
QY 103 GlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArg 122
DB 257 GGGGACAAGAGTGGTACTTCTACGTGCCGAGAGACCGCAAGTACCAGCAACGGCGACCGC 316
QY 123 ProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArg 142
DB 317 CCCAACCGGGTCACGGCGTCCGGCTACTCGAAGGCCACGGCGCCGACCGCAGGCTCTAC 376
QY 143 SerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAla 162
DB 377 GTCGAGGTCAAGCGCGGATCGGGCTCAAGAGACGCTGGTGTCTTACGTGGGGAAGGCG 436
QY 163 ProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisGluThr 182
DB 437 CCCAAGGACTCAGGAGAGCTGGATCATGAACGAGTACCGCTTCTTCCAGCGGAGGCC 496
QY 183 GluLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGlyValGlu 202
DB 497 GACCGCTACCAAG---GAAATTTCTTTGCAAGTGTACAAACGACGAGGATTGAA 553
QY 203 AspHisProSerValProArgSerLeuSerThrArgHisHisAsnHis-AsnSerSerTh 222
DB 554 GACAA-----CTTTCACCTCAGCACCCACACCAAGATCATCTGG 594
QY 222 rSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerSerAsn----- 239
DB 595 CTCAAAGGCTGCAGCAACCATGGAAGAAGACACCGCAACATCGCGCTCGCCCGCAT 654
QY 240 -----HisSerAspAsnLeuAsnAsnAsnAsnAsnAsnIleAs 252
DB 655 GCGCCGCTGTTTCAGCGGTGGGACTCATCAGGTTATCATCAACAG----- 700
QY 252 nAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrTh 272
DB 701 -----CCATACAGTGGAGCAACACCAACCATGCTGGGCATGGC 738
QY 272 rAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIle 287
DB 739 ATCATCGCGGAGCTCGAGCAGCTACTTGGCACCACACAGCGTC 784

Search completed: April 6, 2006, 13:03:38
Job time : 4969 secs

CC modified or altered desirable traits as compared to a reference plant,
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and
 CC flower senescence. Sequence information related to the polynucleotides
 CC and polypeptides can also be used in bioinformatic search methods. The
 CC transgenic plant is useful for growing a progeny plant from a parent
 CC plant. This sequence represents one of the cDNAs of the invention.
 XX
 SQ Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.79e-181 Length: 1444
 Score: 2006.50 Matches: 379
 Percent Similarity: 91.5% Conservative: 0
 Best Local Similarity: 91.5% Mismatches: 0
 Query Match: 98.6% Indels: 35
 DB: 10 Gaps: 1

US-10-780-703-2 (1-379) x ADD30083 (1-1444)

Qy 1 MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
 Db 148 ATGGCAATTGTATCTCCACACACAGCATATCCCATGAGTAAACCAAGTCAACAATAAC 207
 Qy 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlnGluSerHisValGlnAsnGlu 40
 Db 208 GAAAGAAGTATAGAAAGCAATGATCATAGAGCGCCCAAGAGATCATGTCCAAATGAA 267
 Qy 41 AspGluAlaAspAspHisAspHisMetValMetProGlyPheArgPheHisProThr 60
 Db 268 GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 327
 Qy 61 GluGluGluLeuIleGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
 Db 328 GAAGAAGAACTCATAGAGTTTACCTCCCGGAAAGTTGAAGGCAACCGCTTTAATGTA 387
 Qy 81 GluLeuIleThrPheLeuAspLeuThrArgThrArgProThrGluLeuProAlaMetAla 100
 Db 388 GAATCATCATCTTCTCGATCTTATCGTATGATGATGATGATGATGATGATGATGATGATGAT 447
 Qy 101 AlaIleGlyGluLysGluThrPheThrValProArgAspArgLysThrArgAsnGly 120
 Db 448 GCGATAGGAGAAAGAGTGGTACTTCTATGTCACAGAGATCGGAAATATAGAAATGGA 507
 Qy 121 AspArgProAsnArgValThrThrSerGlyThrTrpLysAlaThrGlyAlaAspArgMet 140
 Db 508 GATAGACCGAACCAGTAAACGATTCAGGATATTGGAAGCCACCGGAGCTGATAGGATG 567
 Qy 141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheThrSerGly 160
 Db 568 ATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGGT 627
 Qy 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluThrArgLeuProHisHis 180
 Db 628 AAGCCCTTAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCTTCGCAACCAT 687
 Qy 181 GluThrGluLysThrGlnLysAlaGluLeuSerLeuCysArgValThrLysArgProGly 200
 Db 688 GAAACCGAAGAGTACCAAAAGGCTGAAATATCATTTGTGCGGAGTGTACAAAAGGCGCAGA 747
 Qy 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer 220
 Db 748 GTAGAAGATCATCCATCCGTTACCGTCTCTCTCCACAGACATCATAAACCATTAACATCA 807
 Qy 221 SerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSerSer 240
 Db 808 TCGACATCATCCGTTTACGCTTAGACACACACACACACACATTCATCTCTCTCTATCAT 867
 Qy 241 SerAspAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 260
 Db 868 TCCGACAAACACCTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 927
 Qy 261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280

Db 928 TATTCGGCGACGGCAGCAT 987
 Qy 281 AlaLeuAlaAsnGlnAsnIleThrArgProMetProThrAspThrSerAsnAsnThrLeu 300
 Db 988 GCTCTAGCCCAATCAAAACATATATCGTCCAATGCTTTACGACACAAGCAACACACAT 1047
 Qy 301 IleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValAspAspLeuGln 320
 Db 1048 ATAGTCTCTACGAGAAATCATCAAGACGATGATGAATGCTGCTGCTGCTGCTGCTGCT 1107
 Qy 321 ArgLeuValAsnThrGlnIleSerAspGly----- 330
 Db 1108 AGACTAGTTAACTACCAAAATATCAGATGGAGGTAAACATCAATCAACCAATCTTTCAA 1167
 Qy 330 ----- 330
 Db 1168 GCTCAACAGTTTCATCATCATCAACACAAATAATGCTAACGCAACGCGATTACAAAT 1227
 Qy 331 -----AlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn 345
 Db 1228 GCTCGCGGAGTACAGGACACACGCTTAATGCTTCAAACTCAAGCGGCTTAGCTATGA 1287
 Qy 346 MetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleVal 365
 Db 1288 ATGATTCCTCGAGGACGATTCCAAACAAATGCTTTGTGGATATGTGGAATCCAATAG 1347
 Qy 366 ProAspGlyAsnArgAspHisThrAsnIleProPheLys 379
 Db 1348 CCAGATGGAACACAGAGATCATATCTATTAATTCCTTTTAAG 1389
 RESULT 4
 ADE31520
 ID ADE31520 standard; cDNA; 1444 BP.
 XX ADE31520;
 AC ADE31520;
 DT 29-JAN-2004 (first entry)
 XX
 DE Plant yield related polynucleotide clone G962.
 XX
 KW ds; gene; transcription factor; transgenic plant; salt stress resistance;
 KW osmotic stress resistance; freezing tolerance; drought tolerance;
 KW low humidity tolerance; radiation resistance.
 XX Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 CDS 148..1392
 FT /*tag= a
 FT /product= "transcription factor"
 XX
 FN W02003013228-A2.
 XX
 PD 20-FEB-2003.
 XX
 PF 09-AUG-2002; 2002WO-US025808.
 XX
 PR 09-AUG-2001; 2001US-0310847P.
 PR 19-NOV-2001; 2001US-0336049P.
 PR 11-DEC-2001; 2001US-0338692P.
 PR 14-JUN-2002; 2002US-00171468.
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 FA
 PI Heard JE, Riechmann JL, Creelman RA, Keddle J, Pilgrim ML;
 PI Dubell AN, Jiang C, Ratcliffe O, Pineda O, Yu G, Broun PE;
 XX
 DR WPI; 2003-248222/24.
 DR P-PSDB; ADE31521.
 XX
 PT New plant transcription factor polynucleotides and polypeptides, useful
 in producing transgenic plants with commercially valuable properties,

i.e. modified desirable traits, e.g. salt stress resistance or tolerance to freezing.

The invention relates to a number of isolated cDNA sequences and their encoded proteins which are especially transcription factor related cDNA and proteins. The isolated or recombinant plant transcription factor polynucleotides and polypeptides are useful in producing transgenic plants with commercially valuable properties, i.e. modified or altered desirable traits as compared to a reference plant, e.g. salt stress resistance, osmotic stress resistance, tolerance to freezing, drought, low humidity tolerance, or radiation resistance. Sequence information related to the polynucleotides and polypeptides can also be used in bioinformatic search methods. The transgenic plant is useful for growing a progeny plant from a parent plant. This sequence represents one of the cDNAs of the invention.

Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;

Environment Scores:

NO.:	1,79e-181	Length:	1444
:	2006.50	Matches:	379
Sent Similarity:	91.5%	Conservative:	0
Local Similarity:	91.5%	Mismatches:	0
/ Match:	98.16%	Indels:	35
	10	Gaps:	1

0-780-703-2 (1-379) x ADE31520 (1-1444)

1	MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn	20
148	ATGGCAATTGTATCTCCACAACAAGCATCATTCCTCAGTAAACAAGTCAACAATAC	207
21	GluLysGlyIleGluLeuAspAspHisAtrGlyGlyGlnGluSerHisValGlnAsnGlu	40
208	GAATAAGGTATAGAGACATGATCATAGAGCGGCCAAGAGAGTCATGTCCAAAATGAA	267
41	AspGluAlaAspAspHisAAspMetValMetProGlyPheArgPheHisProThr	60
268	GATGAAGCTGATGATCATGATCATGACATGGTCATGCCCGGATTTAGATTTCATCTACC	327
61	GluGluGluLeuIleGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnVal	80
328	GAAGAAGAATCATAGAGTTTAACTTCGCCGAAGAAGTTGAAGGCCAAACGCTTTAATGTA	387
81	GluLeuIleThrPheLeuAspLeuThrArgThrArgThrAspProTrrpGluLeuProAlaMetAla	100
388	GAACCTCATCTTTCTCTCGATCTTTATCGCTATGATCCTTGGAACTTCTCTGCTATGCG	447
101	AlaIleGlyGluLysGluThrTrpThrPheThrValProArgAspArgLysThrArgAsnGly	120
448	GCGATAGAGAGAAAGAGTGGTATCTTATGTGCCAAGAGATCGGAATATAGAAATGGA	507
121	AspArgProAsnArgValThrThrSerGlyThrTrpLysAlaAlaThrGlyValAspArgMet	140
508	GATAGACCGAACCGAGTAAACATCTCAGAGATTTGGAAAGCCACCAGAGCTGATAGATG	567
141	IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheThrSerGly	160
568	ATCAGATCGGAGACTTCTCGGCCTATCGGATTAAGAAACCCCTAGTATTTCTACTCTGGT	627
161	LysAlaProLysGlyThrArgThrSerTrrpIleMetAsnGluThrArgLeuProHisHis	180
628	AAAGCCCTTAAAGGCATCTCGTACTAGTTGGATCATGAACGAGTATCGCTTCTCCGCACCAT	687
181	GluThrGluLysThrGlnLysAlaGluIleSerLeuCysArgValThrLysArgProGly	200
688	GAATACCGAAGTACCCAAAGGCTGAAATATCATGTGCCGAGTGTCACAAAGGCCAGGA	747
201	ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer	220
748	GTGAAGATCATCCATCGGTACCACTTCTCTCCACAAGATCATCAATCAATCACTCA	807

PA	(JIAN//) JIANG C.
PA	(HEAR//) HEARD J E.
PA	(HAAK//) HAAKE V.
PA	(CREE//) CREELMAN R A.
PA	(RATC//) RATCLIFFE O.
PA	(ADAM//) ADAM L J.
PA	(REUB//) REUBER T L.
PA	(KEDD//) KEDDIE J.
PA	(BROU//) BROWN P E.
PA	(PILG//) PILGRIM M L.
PA	(DUBE//) DUBELL A N.
PA	(PINE//) PINEDA O.
PA	(YUGG//) YU G.
XX	
PI	Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI	Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Brown PE;
PI	Pilgrim ML, Dubell AN, Pineda O, Yu G;
DR	WFI; 2004-132245/13.
DR	P-PSDB; ADI44013.
XX	
PT	New transgenic plant comprising a recombinant polynucleotide of any one
PT	of more than 500 nucleotide sequences, useful in bioinformatic search
PT	methods.
XX	
PS	Disclosure; SEQ ID NO 2475; 435pp; English.
XX	
CC	The invention describes a transgenic plant comprising a recombinant
CC	polynucleotide of any one of more than 500 nucleotide sequences fully
CC	defined in the specification or its complement. The method of the
CC	invention can be used to produced a plant having altered traits such as:
CC	enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
CC	sensitivity; disease resistance; sugar sensing; early or late flowering;
CC	altered flower structure, change in stem bifurcations, altered branching
CC	pattern, reduced apical dominance, reduced trichome density; lack of
CC	trichomes; reduced ectopic trichome development; altered trichome
CC	development; increase in trichome number; altered stem morphology;
CC	increased root growth; increased root hairs; altered seed development;
CC	altered cell proliferation or cell differentiation; rapid development;
CC	premature senescence; increased necrosis; increase in seedling or plant
CC	size; decreased plant size; leaf morphology; seed morphology; seed
CC	biochemistry; increase in root anthocyanins; increase in plant
CC	anthocyanins, or alteration in light response or shade avoidance. The
CC	transgenic plant, polynucleotides and polypeptides are useful in
CC	bioinformatic search methods. This sequence represents a plant
CC	transcription factor related polynucleotide.
XX	
SQ	Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;
	Alignment Scores:
	Pred. No.: 1,79e-181 Length: 1444
	Score: 2006.50 Matches: 379
	Percent Similarity: 91.5% Conservative: 0
	Best Local Similarity: 91.5% Mismatches: 0
	Query Watch: 98.6% Indels: 35
	DB: 12 Gaps: 1
	US-10-780-703-2 (1-379) x ADI44012 (1-1444)
Oy	1 MetAlaIleValSerSerThrThrSeriLeileProMetSerAsnGlnValAsnAsnAen 20
Db	148 ATGGCAATTGTATCTCCACACAAGCATCATTCGCCATGAGTAACCAAGTCAACAATAC 207
Oy	21 GluLysGIyLIeGluaSpAsnAsphSiArgGIyGIyGInGluSerHisValGInAsnGlu 40
Db	208 GAATAAGGTATAGAGAACAATGATCATGAGCGCGCCAAGAGAGTGATGTCACAAATGA 267
Oy	41 AspGIuAlaAspAspHisAspHisAspMetValMetProGLyPheArgPheHisProthr 60
Db	268 GATGAAGCTGATCATGATCATGCATGCATGGTATGCCCGCGGATTTAGATTCCATCTACC 327
Oy	61 GluGIuGIuLeuIIeGluPheTYrLeuArgArgLysValGIuGIyLysArgPheAsnVal 80

Db	328	GAAGAGAACTCATAGAGTTTACCTTCGCCGAAAAGTTGAAGGCCAAACGCTTTAATGTA	387
Qy	81	GlulLeuilethrPheLeuAspLeuTyArGtyrAspProTrpGluLeuProLaMetaLa	100
Db	388	GAACTCATCACTTTCCTCGATCTTTATCGTATGATCCTTGCGAACTTCCTGCTATGGCG	447
Qy	101	AlaileGlyGluLysGluTrpTyPheTyValProArgAspArgLysTyArgAsnGly	120
Db	448	GCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAATATAGAAATGGA	507
Qy	121	AspArgProAsnArgValThrThrSerGlyTyTrpLysAlaThrGlyAlaAspArgMet	140
Db	508	GATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAAGCCACCGAGCTGATAGGATG	567
Qy	141	IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTySerGly	160
Db	568	ATCAGATCGGAGACTTCTCGGCCATTCGGAATTAAAGAAAACCCCTAGTTTTCTACTCTGGT	627
Qy	161	LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyArgLeuProHisHis	180
Db	628	AAAGCCCTTAAAGGCCTCGTACTAGTTGGATCATGAAACGAGTATGTCTTCGACCAT	687
Qy	181	GluThrGluLysTyTrpGlnLysAlaGluIleSerLeuCysArgValTyLysArgProGly	200
Db	688	GAACCCGAGAAGTACCAAAAGGCTGAATATCATTTGTGCCAGTGTACAAAAGGCCAGGA	747
Qy	201	ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer	220
Db	748	GTAGAAGATCATCCATCGGTACACGCTTCTCTCCACAAGACATCATACCACTAACTCA	807
Qy	221	SerThrSerSerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSerSerSerAsnHis	240
Db	808	TCGACATCATCCCGTTTGGCTTAAACAACAACAACACCATTCCTCTCTAATCAT	867
Qy	241	SerAspAsnAsnLeuAsnAsnAsnAsnAlaAsnLeuGluLysLeuSerThrGlu	260
Db	868	TCGCACAACAACCTTAACAACAACAACAACATCAACAATCTCAGAGAAGCTCTCCACCGAA	927
Qy	261	TyrSerGlyAspGlySerThrThrThrThrThrThrAsnSerAsnSerAspValThrIle	280
Db	928	TATTTCGGCGCAGCGCAGACAACAACAACGACCAACAACAGTAACTCTGACGTTACCAT	987
Qy	281	AlaLeuAlaAsnGlnAsnIleTyArgProMetProTyArgThrSerAsnAsnThrLeu	300
Db	988	GCTCTAGCCCAATCAAAACATATATCGTCCAATGCCTTACGACACAAGCAACAACATG	1047
Qy	301	IleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValAspAspLeuGln	320
Db	1048	ATAGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGCCAATGTTGACGATCTTCAA	1107
Qy	321	ArgLeuValAsnTyTrpGlnIleSerAspGly	330
Db	1108	AGACTGTTAACCTACCAAAATATCAGATGGAGGTAAACATCAATCACCNAATCTTCAAATT	1167
Qy	330	-----	330
Db	1168	GCTCAACAGTTTCATCATCATCAACAACAATAATGCTAAACGAAAGCATTTACAATTGGTG	1227
Qy	331	-----AlaThrLeuMetProGlnThrGlnAlaLeuAlaMetAsn	345
Db	1228	GCTCGCGCGACTACAGCGACAACGCTAATGCCCTCAAACTCAAGCGCGTTAGTCTATGAAC	1287
Qy	346	MetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleVal	365
Db	1288	ATGATTCCTGCAGAACGATTCCAACCAATGCTTTGTGGGATATGTGGATCCCAATGTA	1347
Qy	366	ProAspGlyAsnArgAspHisTyThrAsnIleProPheLys	379
Db	1348	CCAGATCGAACAGAGATCACTACTAAATATCTCTTTTAAG	1389

RESULT 6
ADW23849
ID ADW

XX AC ADW23849;
XX DT 10-MAR-2005 (first entry)
XX DE Arabidopsis thaliana LOV1 genomic sequence, SEQ ID 3.
XX KW Flowering; plant; gene; ds.
XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers
XX FT CDS 1..2606
XX FT /*tag= b
XX FT /product= "LOV1 protein"
XX FT exon 1..292
XX FT /*tag= a
XX FT /number= 1
XX FT intron 293..622
XX FT /*tag= c
XX FT /number= 1
XX FT exon 623..892
XX FT /*tag= d
XX FT /number= 2
XX FT intron 893..1923
XX FT /*tag= e
XX FT /number= 2
XX FT exon 1924..2351
XX FT /*tag= f
XX FT /number= 3
XX FT intron 2352..2456
XX FT /*tag= g
XX FT /number= 3
XX FT exon 2457..2606
XX FT /*tag= h
XX FT /number= 4
XX KR2004075252-A.
XX 27-AUG-2004.
XX 20-FEB-2003; 2003KR-00010772.
XX 20-FEB-2003; 2003KR-00010772.
XX (UYSE-) UNIV SEOUL NAT.
XX Ahn JH, Choi EG, Choi YD, Kim YH, Lee JS, Yoo SY;
XX WPI; 2005-054880/06.
XX P-PSDB; ADW23848.
XX New LOV1 polynucleotide encoding a polypeptide controlling the flowering
XX time of plants, i.e., either delaying or inducing early flowering of the
XX plants.
XX Claim 5; SEQ ID NO 3; 25pp; Korean.
XX The present invention relates to the novel gene LOV1 (Long Vegetative
XX phase 1) from Arabidopsis thaliana and its use in controlling flowering
XX time in plants. LOV1 inhibits the flowering stimulating gene AG120. LOV1
XX can be used for manipulating flowering time by over-expressing LOV1 to
XX delay flowering, or inhibiting expression of LOV1 to stimulate flowering.
XX LOV1 is also useful for identifying homologous genes in other plants. The
XX present sequence is the LOV1 genomic sequence.

XX SQ Sequence 2606 BP; 896 A; 509 C; 388 G; 813 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.28e-155 Length: 2606
Score: 1735.00 Matches: 378

Percent Similarity:	43.5%	Conservative:	0
Best Local Similarity:	43.5%	Mismatches:	1
Query Match:	85.3%	Indels:	490
DB:	14	Gaps:	3

US-10-780-703-2 (1-379) x ADW23849 (1-2606)

Qy	1	MetAlaIleValSerSerThrThrSerIleProMetSerAenGlnValAsnAsn	20
Db	1	ATGGCAATTGTATCTCCACACCAAGCATATTCCCATGAGTAACCAAGTCAACAATAAC	60
Qy	21	GlulysGlyIleGluAspAsnAspHisArgGlyGlnGlnGluSerHisValGlnAsnGlu	40
Db	61	GAAGAGGTATAGAGACATGATCATAGAGGGCCCAAGAGAGTCTATGTCCTCAAAATGAA	120
Qy	41	AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr	60
Db	121	GATGAAGCTGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	180
Qy	61	GlulGluGluLeuIleGluPheTyrLeuArgArgGlyValGluGlyLysArgPheAsnVal	80
Db	181	GAAGAGAACTCATAGAGATTTTACCTTCGCCGAAAGTTGAAGCAAAACGGCTTTAATGA	240
Qy	81	GlulLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuPro	97
Db	241	GAATCATCACTTCTCTCGATCTTTATCGTATGATCCTTGGGAACCTTCCTGGTAAATAT	300
Qy	97	-----	97
Db	301	ACATTCACATAAAACACACATAAATCATCTCAAACTATTGGAAATCTTAATTTCTATTCA	360
Qy	97	-----	97
Db	361	TATGTTAAGATCTTCT	420
Qy	97	-----	97
Db	421	ATATATGATACCTACCT	480
Qy	97	-----	97
Db	481	CTTTAAACGCTTCTCCCTCTTTAGTTTCTTTTAAATTAACCTTAATTAACAAACCTACA	540
Qy	97	-----	97
Db	541	TATATATCATAGATATACAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT	600
Qy	98	-----AlaMetAlaAlaIleGlyGluLysGluTyrPheTyr	110
Db	601	TCATAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660
Qy	111	ValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGly	130
Db	661	GTGCCAAGAGATCGGAAATATAGAAATAGAGATAGACCGAACCGAGTAACGACTTCAGGA	720
Qy	131	TyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGly	150
Db	721	TATGGAAAGCCACCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
Qy	151	LeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyr	170
Db	781	TTTAAAGAAACCTAGTTTCTACTCTGTTAAAGCCCTTAAGGCACTCGTACTAGTTGG	840
Qy	171	IleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAla	188
Db	841	ATCATGAACGAGTATCGTCTTCGCCACCATGAACCGAGAAAGTACCAAAAGGT-ATAAAT	899
Qy	188	-----	188
Db	900	TCTACTATAACTCTATATATATATCTTATTCATACATAGATATAACCCCTAGTAGTG	959
Qy	188	-----	188

Db 960 GTGAGCCCTTTAAATTAATCCCTAGACAGTTTGAATTTTCTTTTGTGACTA 1019
Qy 188 ----- 188
Db 1020 GTTTTATTATTATTGGAATTCGATAAGATCAAAATTAATCTTGTGAATGGACTA 1079
Qy 188 ----- 188
Db 1080 AATGTCAGGCGGTTTGGCTTAAATCCAGAAATAATGTTCAATGTCATATGCGTGAATC 1139
Qy 188 ----- 188
Db 1140 TTTAAATTCGTAGACATGCCCCATATGTTATAGTAGAATACATTAATAGATAGATGCATA 1199
Qy 188 ----- 188
Db 1200 CACATATATATAACACACAGATATCACACTCGACATTCATATACCTTAATTTCTGCAGAG 1259
Qy 188 ----- 188
Db 1260 ACATAGTAGTTTCTTACAAATTTATGACATGAATGTTCTGCTCTTCTCCTCACATTAAT 1319
Qy 188 ----- 188
Db 1320 TCATGCTCTTATTAAAGTTACCCACATTTTGTGAATAATTTGGCATATATGAATTAAT 1379
Qy 188 ----- 188
Db 1380 ACCAACATATTATATGCGAACATTTAAATCTATACGAATGATAACGGTTTATGGAGTA 1439
Qy 188 ----- 188
Db 1440 GACCGAAAAATATTATGATATACGGAATGACAAATGGAATAGATATAAATACATTTTGGG 1499
Qy 188 ----- 188
Db 1500 CTCCTTCGACTATATGTCGTCACCATTTGAAACCATAAATTTATAAAATTTTCTATGTA 1559
Qy 188 ----- 188
Db 1560 TATATATGATATTATGATGTATGTCATAGACAGACTAAACACAGGGTTGACATAATTAAT 1619
Qy 188 ----- 188
Db 1620 CTATGATGATGATTGACATTCACATTCGATGTAATAAACTAAATACGCAATTAATA 1679
Qy 188 ----- 188
Db 1680 TATAAAAAATAATAATAATCAATCTTAATTTATATTGTCATTTGTCATATGATAG 1739
Qy 188 ----- 188
Db 1740 TACTCTAAATTTCTTCAAACGCTATCTTTTTTGTCTAAATGCTAACTTTACATAGATT 1799
Qy 188 ----- 188
Db 1800 GTGAATCTTTTCAAACCATATCTTCGATAAATGATATTTTTCATAGATATTGTTAGT 1859
Qy 188 ----- 188
Db 1860 CTATATTTGATAATTTGATATATGATCAAGTCTTAATCAATGTGCTCATGTATAATTA 1919
Qy 189 ----- GluIleSerLeuCysArgValTyrLysArgProGlyValGluAspHisProSer 206
Db 1920 TAGGCTGAATATCATTTGTGCCGAGGTGACAAAGGCCAGGAGTAGAAGATCATCCATCG 1979
Qy 207 ValProArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeu 226
Db 1980 GTACCAAGTCTCTCTCCACAGACATATAACCAATCAATCAATCAATCAATCAATCAAT 2039
Qy 227 AlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSer 246
Db 2040 GCCTTAAGACACAAACACACCAATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2099

Qy 247 AsnAsnAsnIleAsnAsnLeuGluIysLeuSerThrGluTyrSerGlyAspGlySer 266
Db 2100 AACAAACAAACATCAACAATCTCGAAGAGCTCTCCACCGAATATTCGGCGACGCGACG 2159
Qy 267 ThrThrThrThrThrAsnSerAspValThrIleAlaLeuAlaAsnGlnAsn 286
Db 2160 ACAACAACAGCACCACAAACAGTAACCTCGACGTTACCAATGCTCTAGCCAAATCAAAAC 2219
Qy 287 IleTyrArgProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsn 306
Db 2220 ATATATGCTCAATGCGCTTACGACACAGCAACACACATGATGATGATGATGATGATGAT 2279
Qy 307 HisGlnAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGln 326
Db 2280 CATCAAGACGATGATGAACACTGCCATTTGTCGCGATCTTCAAGACACTAGTTAACTACCA 2339
Qy 327 IleSerAspGly----- 330
Db 2340 ATATCAGATGCGAGGTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2399
Qy 331 -----Ala 331
Db 2400 ACTCAACAACAAATGCTTAACGCAACGCAATTAACAATTTGGTGGCTGCGCGCACTACAGCG 2459
Qy 332 ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThr 351
Db 2460 ACAACGCTAATGCTCAAACTCAACGCGCGGTAGTATGAACATGATTTCTTCGAGGAACG 2519
Qy 352 IleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleValProAspGlyAsnArgAsp 371
Db 2520 ATTCCAACAATGCTTTTGGGATATGTGAATCCAAATAGTACCATAGTGAACACAGAGAT 2579
Qy 372 HisTyrThrAsnIleProPheLys 379
Db 2580 CACTATATAATATTCCTTTTAAAG 2603
RESULT 7
ADX17043
ID ADX17043 standard; DNA; 2606 BP.
AC ADX17043;
XX
XX 21-APR-2005 (first entry)
XX Long Vegetative phase 1, LOV1, gene, SEQ ID 3.
DE Long Vegetative phase 1; LOV1; flowering; plant; gene; ds.
XX
XX Arabidopsis thaliana.
OS
XX Key Location/Qualifiers
FH 1. .2606
CDS /tag= b
FT /product= "LOV1"
FT exon 1. .292
FT /tag= a
FT /number= 1
FT intron 293. .622
FT /tag= c
FT /number= 1
FT exon 623. .891
FT /tag= d
FT /number= 2
FT intron 892. .1922
FT /tag= e
FT /number= 2
FT exon 1923. .2352
FT /tag= f
FT /number= 3
FT intron 2353. .2457
FT /tag= g
FT /number= 3

exon 2458..2606
 /*tag= h
 /number= 4
 US2005034194-A1.
 10-FEB-2005.
 19-FEB-2004; 2004US-00780703.
 20-FEB-2003; 2003KR-00010772.
 (UYSE-) UNIV SEOUL NAT IND FOUND.
 Lee JS, Kim YH, Choi EK, Yoo SY, Ahn JH, Choi YD;
 WPI; 2005-054880/15.
 P-PSDB; ADX17042.
 New LOV1 polynucleotide encoding a polypeptide controlling the flowering
 time of plants, i.e., either delaying or inducing early flowering of the
 plants.
 Claim 2; SEQ ID NO 3; 28pp; English.
 The present invention relates to Arabidopsis thaliana Long Vegetative
 phase 1, LOV1, gene (ADX17043), coding sequence (ADX17041) and protein
 (ADX17042). LOV1 controls the flowering time of plants by repressing the
 activity of flowering-promoting gene AGL20, and so is useful for
 controlling flowering: the flowering time of plants can be delayed by
 LOV1 gene overexpression, or early flowering can be induced by
 suppressing the expression of the LOV1 gene. The LOV1 gene was isolated
 from Arabidopsis thaliana by an activation tagging screen.
 SQ Sequence 2606 BP; 896 A; 509 C; 388 G; 813 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3, 28e-155 Length: 2606
 Score: 1735.00 Matches: 378
 Percent Similarity: 43.5% Conservative: 0
 Best Local Similarity: 43.5% Mismatches: 1
 Query Match: 85.3% Indels: 490
 DB: 14 Gaps: 3
 US-10-780-703-2 (1-379) x ADX17043 (1-2606)
 Qy 1 MetAlaIleValSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
 Db 1 ATGGCAATTGTATCTCCACAAAGCATCATTCCTCATGAGTAACCAAGTCAACAATAAC 60
 Qy 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGlnSerHisValGlnAsnGlu 40
 Db 61 GAAAGAGGTATAGAGCAAAATGATCATAGAGCGGCCCAAGAGAGTGTCTCCAAATGAA 120
 Qy 41 AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr 60
 Db 121 GATGAAGCTGATGATCATGATCATGACATGATGATGATGATGATGATGATGATGATGAT 180
 Qy 61 GluGluGluLeuIleGluPheTyrluArgArgLysValGluGlyLysArgPheAsnVal 80
 Db 181 GAAGAAGAACTCATAGAGTTTTTACCTTCGCCGAAAGTTGAAGCAACGCTTTAATGTA 240
 Qy 81 GluLeuIleThrPheLeuAspLeuTyrluArgTyrluArgProThrGluLeuPro 97
 Db 241 GAACTCATCTTCTCTCGATCTTATCGTATGATCTCTGGAACCTTCTCTGGTAATAT 300
 Qy 97 ----- 97
 Db 301 ACATTACATAAACACACATAAATCATCTCAAACTATTGGAATCTTAATTTCTATTCA 360
 Qy 97 ----- 97
 Db 361 TATGTTAAGATCTTCTCTCTCTATACATCTTCTCTCTCTATTTCTTTTTTTTAACT 420
 Qy 97 ----- 97
 Db 421 ATATATGTACTACCTCTCTTATGAAGTATTACTATGTCTGATCGTTAACAATTTCTCAATAT 480
 Qy 97 ----- 97
 Db 481 CTTTAAACGCTTCTCTCTCTTCTTTCTTTTAAATTAACCTTAATTAACAACACTACA 540
 Qy 97 ----- 97
 Db 541 TATATATCATAGATATACAAATATATGTATATGTTTTCATTAATTAAGTTATGTATGTTTAA 600
 Qy 98 ----- AlaMetAlaAlaIleGlyGluLysGluTyrluPheTyrlu 110
 Db 601 TCATAGATATATGATATGCGCTATGCGCGCATAGAGAGAAAGAGTGTCTACTTCTAT 660
 Qy 111 ValProArgAspArgLysTyrluArgAsnGlyAspArgProAsnArgValThrThrSerGly 130
 Db 661 GTGCCAAGAGATCGGAAATATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGGA 720
 Qy 131 TyrluPheAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGly 150
 Db 721 TATTGGAAAGCCACCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 Qy 151 LeuLysLysThrLeuValPheTyrluSerGlyLysAlaProLysGlyThrArgThrSerTrp 170
 Db 781 TTAAGAAACCTTAGTTTCTACTCTGTTAAGCCCTTAAGGCACTCGTACTAGTTGG 840
 Qy 171 IleMetAsnGluTyrluArgLeuProHisGluThrGluLysTyrluGlnLysAla 188
 Db 841 ATCATGAACGAGTATCGTCTTCGACCATGAAACCGAGAGTACCAAAAGGT-ATAAAT 899
 Qy 188 ----- 188
 Db 900 TCTACTATAACTCTATATATATCTATTCATACATACATAGATATAACCTAGTAGTG 959
 Qy 188 ----- 188
 Db 960 GTGAGGCTTTAAATTTGAATTAATCCCTAGACAGTTTGTGATTTTCTTTTGTGACTA 1019
 Qy 188 ----- 188
 Db 1020 GTTTTATTTATTTTGTGAATTTGATTCGATGAATCAAAAATACTTGTGAATGGACTA 1079
 Qy 188 ----- 188
 Db 1080 AATGTACAGCGCGCTTTGCGCTTAAATCCAGAAAAATGTTTCATGTCATATCGTGAAC 1139
 Qy 188 ----- 188
 Db 1140 TTTAAATTTGTAGACATGGCCCATATGTTATAGTAGAATACATTAATAGATAGATG 1199
 Qy 188 ----- 188
 Db 1200 CACATATATATAACACACAAGTATCACACTCGCAATTCATATACCTTAATTTCTGCAG 1259
 Qy 188 ----- 188
 Db 1260 ACATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCTCTCTCTCTCCTCACATTA 1319
 Qy 188 ----- 188
 Db 1320 TCATGCTCTCTATTAAGTTTACCAACATTTTGAATTAATTTGGCATATATGAATTA 1379
 Qy 188 ----- 188
 Db 1380 ACCAACATATTTATATGCGAACATTTAAATCTATACGAATGATAACGGTTTATGGAGTA 1439
 Qy 188 ----- 188
 Db 1440 GACCGAAAAAATATATGTTATACGAAAAATGCAATGGATAGATAAATACATTTTTTTGGG 1499


```
QY 188 ----- 188
Db 1500 CTCCTTCGACTTATATGTCGTCCACCATTTGAAACCATATAAATTTTATGTA 1559
QY 188 ----- 188
Db 1560 TATATATGATATTATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
QY 188 ----- 188
Db 1620 CTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
QY 188 ----- 188
Db 1680 TATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1739
QY 188 ----- 188
Db 1740 TACTCTAAATTTCTTAAACGTGCTCTCTTTTGTCTAATGCTAACTTTACATAGTTT 1799
QY 188 ----- 188
Db 1800 GTGAATCTCTTTTCAAAACCATATCTTCGATAAATGATATTTTTCATAGATTTGTTAGT 1859
QY 188 ----- 188
Db 1860 CTATATTTGATTAATTTGATATGATATGATCAAGTCTCTAATCAATGCTCATGATAAATTA 1919
QY 188 ----- 188
Db 1920 TAGGCTGAATATCATTTGTCGAGTGATCAAAAGGCGAGGAGTAGAGATCATCCATCG 1979
QY 207 ValProArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeu 226
Db 1980 GTACCAGGTTCTCTCCCAAGACATCATCAACCAATCACTCATGACATCATCCCGTTTA 2039
QY 227 AlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSer 246
Db 2040 GCCTTAAGACAAACAAACACCATTCCTCTCTTAATCATTCGACAAACCTTTAAC 2099
QY 247 AsnAsnAsnAsnIleAsnAsnLeuGluLeuSerThrGluTyrSerGlyAspGlySer 266
Db 2100 AACACACAAACATCAACATCTCGAGAGCTCTCCACGATATTCGCGCAGCGCAGC 2159
QY 267 ThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 286
Db 2160 ACAACAACACGACCAACAAAGTAACTCTGACGTTACCATTTGCTTAGCCAAATCAAAAC 2219
QY 287 IleTyrArgProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsn 306
Db 2220 ATATATCGTCCAAATGCTTTACGACAAACGACAAACACATTCATGCTCTACGAGAAAT 2279
QY 307 HisGlnAspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGln 326
Db 2280 CATCAGACGATGATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2339
QY 327 IleSerAspGly----- 330
Db 2340 ATATCAGATGGAGGTAAACATCAATCAACCAATCTTTTCAAAATGCTCAACAGTTTCATCAT 2399
QY 331 -----Ala 331
Db 2400 ACTCAACAAACAAATGCTAAACGCAACGCAATTCATTTGCTGCGCGACTACAGCG 2459
QY 332 ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThr 351
Db 2460 ACAACGCTAATGCTCTCAAACTCAAGCGGCTTAGCTATGAAATGATTTCTTCGAGGAGC 2519
QY 352 IleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleValProAspGlyAsnArgAsp 371
Db 2520 ATTCCAAACAATGCTTTGTGGGATATGTGGAATCCAAATAGTACCAGATGGAACAGAGAT 2579
QY 372 HisTyrThrAsnIleProPhelys 379
```

```
Db 2580 CACTATACTATATATCTTTTAAG 2603
RESULT 8
ADA70597
ID ADA70597 standard; DNA; 1212 BP.
AC ADA70597;
XX 20-NOV-2003 (first entry)
DT Rice gene, SEQ ID 3920.
DE Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX Oryza sativa.
XX WO2003000898-A1.
PN 03-JAN-2003.
PD 22-JUN-2001; 2001WO-IB001105.
PF 22-JUN-2001; 2001WO-IB001105.
PR 22-JUN-2001; 2001WO-IB001105.
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX Claim 6; SEQ ID NO 3920; 899pp; English.
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
SQ Sequence 1212 BP; 245 A; 421 C; 361 G; 185 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.33e-64 Length: 1212
Score: 781.00 Matches: 185
Percent Similarity: 57.1% Conservative: 47
Best Local Similarity: 45.6% Mismatches: 94
Query Match: 38.4% Indels: 81
Db: 8 Gaps: 11

US-10-780-703-2 (1-379) x ADA70597 (1-1212)
QY 52 MetProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeuArgArg 71
Db 1 ATGCCCGGTTTCAGGTTCCACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
QY 72 LysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyr 91
Db 61 AAGGTGGAGGGCAACGCGTTCAACGTCGAGTCATCCTTCCTCGACTTACCGCTAC 120
QY 92 AspProTyrGluLeuPro-----Ala 98
```

Db 121 GACCCCTGGAGCTCCCGGAATTAAGCATCTAAATAAAGAGCTCGATGGGGTGCAGCA 180
Qy 99 MetAlaAlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArg 118
Db 181 ATGGCGGGATAGGGGAGAGAGAGTGGTCTTCTACGTGCGCTCGGGACAGAGTACAGG 240
Qy 119 AenGlyAspArgProAsnArgValThrThrSerGlyTyrTyrPheAlaThrGlyAlaAsp 138
Db 241 AACGGGACCGCGCGAACCAGCGGTACGGGTCTGGAAGCGGACGGGGCGCGAC 300
Qy 139 ArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyr 158
Db 301 CGGATGATCCGAGCGGAGAACACCGCCCATCGGGCTCAAGAACAGCGTGTCTCTTAC 360
Qy 159 SerGlyLysAlaProLysGlyThrArgThrSerTyrIleMetAsnGluTyrArgLeuPro 178
Db 361 TCCGGCAGGCCCCCAGGGCGTCCGACGAGCTGATCATGAACAGATTCGCCCTTCCC 420
Qy 179 HisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeuCyseArgValTyrLysArg 198
Db 421 CCGCGCGACCGCGCTACCAAGAGCTGAATCTCACTATGCGCGTGTACAAAGCGC 480
Qy 199 ProGlyValGluAsp---HisProSerVal-ProArg-----Se 210
Db 481 ACCGCGATCGACGCGCGCATGGCCAGGTATCCACGGCGCGTCTCGCGCGCACTCCGCG 540
Qy 210 rLeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgG1 230
Db 541 GCGCGCGCGCGCGCATCGGTACAGAGAACAGCAAGAGGCTCTCTGTCAGCTCCACGCGG 600
Qy 230 n-----GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 601 ACGGCGCTCCCACTCCGCTCCAGCTCCACTCTCTCAGCAGCGAGTGTCAGCTCGCGGCC 660
Qy 247 nAsnAsnAsnLeuAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerTh 267
Db 661 GCCATCGTCGCGGACCGCGCCCATGGTGGCGCACAGGCGACCGCTCGCGCGCGG----- 715
Qy 267 rThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnI1 287
Db 716 -ACCATCAGCAGCAGCAGCTCCACGCTGCCAAGCCTTCGCGGGTACTCGCAGAAAT 774
Qy 287 eTyrArg-----ProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrAr 305
Db 775 TCCTCGATGGCTTCGCTCGCGCGCGCGCATCAGCAGCAGCAATTTACCAAGACTTCGCG 834
Qy 305 g----- 305
Db 835 GCGGCATTGTACAGCAGTACTCCAAAGAACAGCAGCGGTGCTTCGCGCTCCACGCTACTCG 894
Qy 306 -----AsnHisGlnAspAspAspGluThrAlaIleValAspAspLe 319
Db 895 CTGCTCAACCTCGTCAACGCGGCATCCATG-GGCAGCTCCGCGCTGCCATGTGATGAGCT 953
Qy 319 uGlnArgLeuValAsnTyr-----GlnIleSerAspGlyAlaThrThr----- 333
Db 954 GAGCTCGTGTGGCCACGCGCACCGCTCTCTACATTAACCCCGCGCGCGCGCACCA 1013
Qy 334 -----LeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIlePr 348
Db 1014 CTATAGTCAATTCTTCACTCTTACGACGCGCGTCTCGCATCAACCAACGCGCGCGCC 1073
Qy 348 aAlaGly-----ThrIlePro-----As 354
Db 1074 GCTCGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCTC 1133
Qy 354 nAsnAlaLeuTyrAspMetTyrAsnProIleValProAspGlyAsnArgAspHisTyrTh 374
Db 1134 CGACAGGATCTGGGACTGGAAATAATCCCATCCCGAGGCGCGGAGTAGGGATTACAGCAC 1193
Qy 374 rAsnIleProPheLys 379
Db 1194 CAGTACTGGATTCAAG 1209

RESULT 9

ACL26812
ID ACL26812 standard; cDNA; 1212 BP.

XX ACL26812;

AC 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polynucleotide SEQ ID NO:768.

XX ss; abiotic stress tolerance; transgenic plant; cereal;
KW agriculture.

XX Oryza sativa.

XX WO2003008540-A2.

XX 30-JAN-2003.

XX 21-JUN-2002; 2002WO-US019668.

XX 22-JUN-2001; 2001US-0300112P.

XX 24-AUG-2001; 2001US-0314662P.

XX 26-SEP-2001; 2001US-0325277P.

XX 21-NOV-2001; 2001US-0332132P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
PI Moughamer T, Provart N, Rickes D, Zhu T;

XX WPI; 2003-248011/24.

XX New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.

XX Claim 1; SEQ ID NO 768; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

XX SQ Sequence 1212 BP; 245 A; 421 C; 361 G; 185 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.33e-64 Length: 1212
Score: 781.00 Matches: 185
Percent Similarity: 57.1% Conservative: 47
Best Local Similarity: 45.6% Mismatches: 94
Query Match: 38.4% Indels: 81
DB: Gaps: 11

US-10-780-703-2 (1-379) x ACL26812 (1-1212)

Qy 52 MetProGlyPheArgPheHisProThrGluGluLeuIleGluPheTyrLeuArgArg 71

Db 1 ATGCCCCGGGTTCAGTTTCCACCCCGAGGAGGAGTGTATCGATCTACCTCGCGCGG 60

Qy 72 LysValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyr 91

Db 61 AAGGTGGAGGCAAGCGCTTCAACGTCGAGCTCATCACCTTCTCGACCTCTACCGCTAC 120
Qy 92 AspProTTPGlulLeuPro-----Ala 98
Db 121 GACCCCTGGAGCTCCCGGAATTGAGCATCTAATTAAACAGCTCGATGGGGTGCAGCA 180
Qy 99 MetAlaAileGlyLulysGluTptTyrPheTyrValProArgAspArgLysTyrArg 118
Db 181 ATGGCGGCATAGGAGAGAGAGTGTCTTCTACGTGCTCGGACAGGAGTACAGG 240
Qy 119 AsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrrLysAlaThrGlyAlaAsp 138
Db 241 AACGGGACCGGCGGAACCGGTGACCGGTGCGGTACTGGAAGCGACGCGGGCCGAC 300
Qy 139 ArgMetIleArgSerGluThrSerArgProIleGlyLulysThrLeuValPheTyr 158
Db 301 CGGATGATCGAGCCGAGACAACACCGCCATCGGGCTCAAGAAGACGCTTGTCTTCTAC 360
Qy 159 SerGlyLysAlaProLysGlyThrArgThrSerTrrPileMetAsnGluTyrArgLeuPro 178
Db 361 TCGGGACAGCCCGCCAGGGGTCCCGACGAGCTGGATCATGAACGATCGCTTCC 420
Qy 179 HisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeuCysArgValTrrLysArg 198
Db 421 CCGCGCACACCGCGCTACCAACAGACTGAAATCTCACTATGCCGCGGTGTACAAGCGC 480
Qy 199 ProGlyValGluAsp---HisProSerVal-ProArg-----Se 210
Db 481 ACCGGCATCGACGACGCCATGCCAGGTATCCACGGCGCGGTGTCGGCGCACTCCCGC 540
Qy 210 rLeuSerThrArgHisHisAsnHisAsnSerThrSerSerArgLeuAlaLeuArgL 230
Db 541 GCGCGCGCGCGCACCGGTACAGGACAAAGCAAGCTCTCTGTCGAGCTCCAGCGCG 600
Qy 230 n-----GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 601 ACGCGCGCTCCAACTCCGTCGAAGTCCACCTCTCAGCAGCGAGTGACGTCGCGCGCC 660
Qy 247 nAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerTh 267
Db 661 GCCATCGTCAGCGACACCGCCCATGTGTGGCGCACAAAGCCGCGTCGCGCGCGC----- 715
Qy 267 rThrThrThrThrThrThrAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnI 287
Db 716 -ACCATCAGCAGCAGCAGCAGCTCCAGCTGCCAGCTTGGCGCGCTACTCTGAGAT 774
Qy 287 eTyrArg-----ProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrAr 305
Db 775 TCCTGATGGCTTCGCTCGCGCGCGCGATCAGCAGCAATTTTCAGCAAGACTTCGCG 834
Qy 305 g----- 305
Db 835 GCGGCATGTACAGCAGTACTCAAGAACACAGCGGTGCTTCCGCTCCACGTACTCG 894
Qy 306 -----AsnHisGlnAspAspAspGluThrAlaIleValAspAspLe 319
Db 895 CTGCTCAACCTCGTCAACCGCGCATCCATG-GGCAGCTCCGCGCTGCCATGTAGCT 953
Qy 319 uGlnArgLeuValAsnTyr-----GlnIleSerAspGlyAlaThrThr----- 333
Db 954 GAGCTCGTGTGGCCACCGCACCGCGCTCTACATTAAACCGCGCGCGCGCGCACAA 1013
Qy 334 -----LeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIlePr 348
Db 1014 CTATAGTCAATTCTTTCACCTGCTACGACGCGCTGCTGCATCAACCAACCGCGCGCG 1073
Qy 348 oAlaGly-----ThrIlePro-----As 354
Db 1074 GTCGGGACGACGACGCGCGCGCGCGCTGCGCGATGTCGCTCGCGCGCTTCTC 1133
Qy 354 nAsnAlaLeuTrrAspMetTrrAsnProIleValProAspGlyAsnArgAspHisTyrTh 374
Db 1134 CGACAGGATCTGGGATCTGGATATTCGATCCCGAGCGCGGAGGTAGGGATTACAGCAC 1193

Qy 374 rAsnIleProPheLys 379
Db 1194 CAGTACTGGATTCAAG 1209

RESULT 10

ADA71117

ID ADA71117 standard; DNA; 1250 BP.

XX

AC ADA71117;

XX

DT 20-NOV-2003 (first entry)

XX

DE Rice gene, SEQ ID 4440.

XX

KW Plant; bacterial infection; fungal infection; viral infection; rice;

KW gene; ds.

XX

OS Oryza sativa.

XX

PN WO2003000898-A1.

XX

PD 03-JAN-2003.

XX

PF 22-JUN-2001; 2001WO-IB001105.

XX

PR 22-JUN-2001; 2001WO-IB001105.

XX

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX

DR WPI; 2003-175290/17.

XX

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 6; SEQ ID NO 4440; 899pp; English.

XX

The present invention relates to a method (M1) for identifying genes
involved in plant resistance or response to pathogenic infection. M1
comprises identifying a gene whose expression is significantly altered in
the incompatible interaction of plant gene expression relative to
expression of the gene in an uninfected plant, in a mutant plant that
does not express a gene associated with response to pathogenic infection,
or in a corresponding incompatible or compatible interaction. (M1) is
useful for conferring resistance to resistance or tolerance to a plant to
bacterial, fungal or viral infection. The present sequence was used to
illustrate the invention.

XX

SQ Sequence 1250 BP; 291 A; 388 C; 339 G; 230 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	1,748-62	Length:	1250
Score:	759.00	Matches:	174
Percent Similarity:	50.8%	Conservative:	45
Best Local Similarity:	40.4%	Mismatches:	72
Query Match:	37.3%	Indels:	140
DB:	8	Gaps:	14

US-10-780-703-2 (1-379) x ADA71117 (1-1250)

Qy 50 MetValMetProGlyPheArgPheHisProThrGluGluLeuLeuPheTyrLeu 69

Db 1 ATGGTAATGCCGGGTTCCGGTTCCACCCGAGGAGGAGCTCATCGATTCTACCTC 60

Qy 70 ArgArgLysValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyr 89

Db 61 CGCGGAAAGGTGGAGGGGAAACGGTTCAACATCGAGCTCATCGCTTCTCGACCTTAC 120

```
Qy 90 ArgTyrAspProTyrGluLeuProAlaMetAlaIleGlyGluLeuTyrPhe 109
Db 121 CGCTATGACCGTGGGATCTTCCGCTCTGGCTCAATTGGGCAAGAGTGGTCTTC 180
Qy 110 TyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSer 129
Db 181 TATGTGCCAAGGAGCGCAAGTACCAAGACCGGATCGGCCCAACCGAGTACGCCATCG 240
Qy 130 GlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIle 149
Db 241 GGGTACTGGAGGCCACAGGGCGGATAGTAGTGGTGAAGGTGAGGTGACCGTCTATC 300
Qy 150 GlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSer 169
Db 301 GGCCTCAAGAGACGCTCGTCTTCTACGTGCGCAAGGCGCCCAAGGGGCTTCGCAGCAGC 360
Qy 170 TrpIleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLys 187
Db 361 TGGATCATGACGAGTATGCTCTCCCTCATGGCGACGCTGACCGCTACCAAAAGTAGCT 420
Qy 187 ----- 187
Db 421 TTTGATATAGTACCGGAGGCGAGCAAGCATATGCTACTGCTATGCTATTGGCAGCTCT 480
Qy 188 -----AlaGlu 189
Db 481 TGTCTCTCAGCTGGTGATGAATTAAGTACACACCATCGGAATTTGTATGAGGAA 540
Qy 190 IleSerLeuCysArgValTyrLysArgProGlyValGluAspHis 204
Db 541 ATTTGCTGTGTGAGTGTACAAAGCCCGAGGATCGAAGCAACTTCCACTVACCGGC 600
Qy 205 -----ProSerValProArgSerLeuSerThrArgHis 215
Db 601 ACAACCACTAAATCTCCGGTCAAGGCGGCGGCGATGGGCAAGCAAGCAGCGCGMC 660
Qy 216 HisAsnHisAsnSerThrSerArgLeuAlaLeuArgGlnGlnGln 232
Db 661 CGAACCGGAGTCTACACGGCGGCGCTGGGCGGCCCTCGCCCGCATGTTTCGAC 720
Qy 233 -----HisHisSerSer 241
Db 721 GCGGCGGCGGCGGCGCACCTCATCTCCGCTCATGGGAAGGCAACATACACGCTCA 780
Qy 242 AsnAsnAsnLeuAsnAsnAsnAsn 251
Db 781 CACACCGCGCGGAGGAGCAAGCCCATCGCCATGTCACGTGCGGCGGCGCGCGCTG 840
Qy 252 -----AsnAsnLeuGluLys 256
Db 841 CATCTCCGCGCGCGGCGATGTTCCGTCGACGGCGTGGTGGCTCGCTGAGTCCACC 900
Qy 257 LeuSerThrGluTyrSerGlyAsp 269
Db 901 ACGTCGACGAGGAGGAGCGACGCGCTCTTCCACCTCAAGGCGGCGGCGACGACG 960
Qy 270 ThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArg 289
Db 961 AACACGAGC -----ACGATGTCATCATCCATCATCATCATCAG 996
Qy 290 ProMetProTyrAspThr -----SerAsnAsnThrLeuIleValSerThr 304
Db 997 CAGTGTCTGCTTCACTCATGCTCTGCTCAACACCAACTGTCAGCAATGGCG ----- 1050
Qy 305 ArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGlnArgLeuVal --- 323
Db 1051 -----ACCATCCCATCGAGGCTGAGTAGGCGGATGGA 1086
Qy 324 -----AsnTyrGlnIleSerAspGlyAlaThrThrLeu 334
Db 1087 TCTTACAGCAACATCCACCATCAACCAAACTCTACTACAACTCATCACTCAACTGCTTTG 1146
Qy 335 MetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsn 354
```

```
Db 1147 CCGCCATCGCAACAAGCCCTCTT -----CTTCCATTGCTAGCATG----- 1188
Qy 355 AsnAlaLeuTrpAspMetTrpAsnProIleVal 365
Db 1189 GAGAAGATTGGGAT---TGGAACCTCTCTCTA 1218
RESULT 11
ADA70311
ID ADA70311 standard; DNA; 549 BP.
XX
AC ADA70311;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 3634.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
PD
PP 03-JAN-2003.
PP 22-JUN-2001; 2001WO-IB001105.
PP 22-JUN-2001; 2001WO-IB001105.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
DR WPI; 2003-175290/17.
PT
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
PS Claim 6; SEQ ID NO 3634; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
SQ Sequence 549 BP; 103 A; 174 C; 194 G; 78 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.05e-56 Length: 549
Score: 688.50 Matches: 125
Percent Similarity: 88.8% Conservative: 10
Best Local Similarity: 82.2% Mismatches: 12
Query Match: 33.8% Indels: 5
DB: 8 Gaps: 1
US-10-780-703-2 (1-379) x ADA70311 (1-549)
Qy 32 GlyGlnGluSerHisValGlnAsnGluAspGluAla-----AspAspHis 46
Db 49 GGGAGAGAGCGCGCGCGCGATCAGGAGGAGCGCGCGCGCTGTCGCGGCGGACTCGCAC 108
Qy 47 AspHisAspMetValMetProGlyPheArgPheHisProThrGluGluLeuIleGlu 66
-----
```

109 GAGAACGACCTGTGTATGTCGGGGTTCGGTTCCACCCGAGGAGGAGGAGCTGATCGAG 168
 67 PheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeu 86
 169 TTCCTACCTCCGGCGGAGGTGGAGGCGCGGTTCACGTCGAGCTCATCCTTCCTC 228
 87 AspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAlaAlaGlyGluGlyGlu 106
 229 GACCTTACCGTTCGACCCATCGGAGCTCCCGCGCATGCGGTGATAGGGAGAGAG 288
 107 TrpTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgVal 126
 289 TGGTTCTTCTACGTGCGCGGACCGCAAGTACCGGACCGCGGACCGCGGACCGG 348
 127 ThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSer 146
 349 ACGGCGTCGGGTACTGGAAGCGCGCGGCGCGCGGATGATCCGCGCGGAGACAGC 408
 147 ArgProIleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThr 166
 409 CGCCCATCGGCTCAAGAGAGCGCTCTCTTACTCCGGCAGGCGCCCGAGGGCGTC 468
 167 ArgThrSerTrpIleMetAsnGluTyrArgLeuPro 178
 469 CGCAGCAGCTGGATCATGAACGAGTACCGCTTCCCC 504

RESULT 12

ADW18167

ID ADW18167 standard; cDNA; 1563 BP.

XX AC ADW18167;

XX DT 24-MAR-2005 (first entry)

XX DE Pinus radiata transcription factor cDNA NAC family Seq 1946.

XX KW gene; ss; plant; transcription; gene regulation; gene expression;

XX KW transgenic plant; drought resistance; disease resistance; salt tolerance;

XX KW cold tolerance; freezing tolerance; flowering; flavor enhancer;

XX KW flower color.

XX OS Pinus radiata.

XX PN WO2005001050-A2.

XX XX 06-JAN-2005.

XX XX 07-JUN-2004; 2004WO-US017965.

XX XX 06-JUN-2003; 2003US-0476189P.

XX XX (ARBO-) ARBORGEN LLC.

XX XX Blokeberg LN, Bryant C, Connert MB, Emerson SV, Frost MJ;

XX XX Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;

XX XX Phillips J, Puthgae S, Veerakone S, Westwood C, Gause K, Wood M;

XX XX WPI; 2005-075542/08.

XX XX P-PSDB; ADW18488.

XX XX Claim 3; SEQ ID NO 1946; 1265pp; English.

XX CC This invention relates to novel isolated plant nucleic acid molecules, or

XX CC variants thereof, that encode transcription factors. Specifically, it

XX CC refers to transcription factor proteins that are capable of binding to

XX CC DNA in order to regulate gene transcription and gene expression in a

XX CC plants, in particular Eucalyptus grandis and Pinus radiata. The present

XX CC invention describes DNA constructs containing DNA encoding a

XX CC transcription factor that regulates the promoter, which is operably

CC linked to the desired nucleic acid to be expressed. It further provides
 CC transgenic plants expressing a transcription factor that confers a trait
 CC to the plant such as increased drought, salt or disease tolerance, height
 CC change, enhanced cold/ frost tolerance, enhanced color, health and
 CC nutritional characteristics, as well as improved taste, starch
 CC composition, flower longevity and germination, amongst others.
 CC Accordingly, such plants that are successfully transformed with a DNA
 CC construct can be characterized by a difference in flower color, petal or
 CC leaf shape and size, aroma or plant height. This polynucleotide is a
 CC plant transcription factor cDNA sequence of the invention.

XX SQ Sequence 1563 BP; 501 A; 280 C; 369 G; 413 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,41e-55	Length:	1563
Score:	688.00	Matches:	160
Percent Similarity:	53.6%	Conservative:	42
Best Local Similarity:	42.4%	Mismatches:	119
Query Match:	33.8%	Indels:	56
DB:	14	Gaps:	9

US-10-780-703-2 (1-379) x ADW18167 (1-1563)

QY	38	GlnAsnGluAspGluAlaAspAsp	-----HisAspHisAspMetValMetPro	53
DB	145	CGCAAGAGAGTCAAGGCGAGATGAAGAGAGCTTCAAGCAATCAGGATGTGTCATCGCG	204	
QY	54	GlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTyrLeuArgArgLysVal	73	
DB	205	GGTTTCAGATTCCATCCCTACGGAAGAAGAAATGTTGGAGTTTTATCTCGCGCAAGAAGTG	264	
QY	74	GluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspPro	93	
DB	265	GAGGGGAAGCACTTCAACATTTAGAGTTATTACACCGTGGATCTTTATAGATACGATCCC	324	
QY	94	TrpGluLeuProAlaMetAlaAlaGlyGluLysGluTyrPheTyrValProArg	113	
DB	325	TGGGAGCTTCCAGCTTTCGCTTCAATGGAGAGAAGAGTTGTTCTTTTACGTATACGCA	384	
QY	114	AspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSerGlyTyrTrpLys	133	
DB	385	GATAGAAAAGTAGGAATGGAGATCGACCTAACAGATTACAAAGTCTGGGTATTGGAAA	444	
QY	134	AlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLys	153	
DB	445	GCTACAGTGTGACCGGATGGTTCCAGTGAAGTCTCTGGTTGCATCGCTTGAAGAAA	504	
QY	154	ThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsn	173	
DB	505	ACGCTTGTGTTCTATACAGGGAAGCGCCCAAGGAAAGCAAGCAAGCTGGATCATGACC	564	
QY	174	GluTyrArgLeuProHisGluThrGluLysTyrGlnLysAlaGluLeuSerLeuCys	193	
DB	565	GAATATAGTGTGCCCACTCGAAATTCGAAATTCGAAAGAACCAAGCAATCTCCCTTTC	624	
QY	194	ArgValTyrLysArgProGlyValGluAspHisProSerValProArgSerLeuSerThr	213	
DB	625	CCGTTTCGAAAAGGTGAGCA-----AAATCCGGAAGGAGGTTTGGAGCA	672	
QY	214	ArgHisHisAsnHisAsnSerSerThrSerSerArgLeu-----	226	
DB	673	AAGGAACATGTTAAGGGAACCGGAACTGATAGGCAACACTTATAAAGATAATGTTGA	732	
QY	227	---AlaLeuArgGlnGlnHisHisSerSerSerSerAsnHisSerAspAsnLeu	245	
DB	733	CCTTGTATTACATTTGAAATCGCAAGATGATGTCGATGTCGATGTCGATGTCGATG	792	
QY	246	AsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGly	265	
DB	793	GCATCCACACATTAAGAACACCGTTCGAGTCTTGAATTTGGAGATCAAGCA---GGA	849	
QY	266	SerThrThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAsnGln	285	

QY 300 LeuileValSerThrArgAsnHisGlnAspAspGluThrAlaileValAspAspLeu 319
 DB 880 CTGCTCTTCGACCGCAGTTGGCGCCTGACGATCAGATGATTTGGTGCAT--- 936
 QY 320 GlnArgLeuValAsnTyrGlnIleSerAspGlyValaThrThrLeuMetProGlnThrGln 339
 DB 937 -----TCATCTGTAATGCAATGGAGGAGGCATCAGCAACAGAGCCAGATCCA 987
 QY 340 AlaAlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuTyrAsp 359
 DB 988 TCAGCAAGAAAG-----CCTGGCAGTGGCTTCAGCATGAACACAGTGAATGGGAT 1035
 QY 360 -----MetTTPAsnProIleValProAspGlyAsnArgAspHisTyr 373
 DB 1036 GGTGTGGCAAGAAATCAACTCCGTTGATTTAGGAGCAGATCTTCAGAGAGATGGAGG 1095
 QY 374 ThrAsnIleProPhelys 379
 DB 1096 AGCAGCATACCTTGGGAA 1113
 RESULT 14
 ID ADA70350 standard; DNA; 950 BP.
 AC ADA70350;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice gene, SEQ ID 3673.
 KW Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX
 OS Oryza sativa.
 XX
 PN WO2003000898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-IB001105.
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX
 DR WPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 XX gene expression.
 PS Claim 6; SEQ ID NO 3673; 899pp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC the expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 950 BP; 186 A; 319 C; 312 G; 133 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,31e-41 Length: 950
 Score: 536.00 Matches: 107

Percent Similarity: 60.4% Conservative: 44
 Best Local Similarity: 42.8% Mismatches: 81
 Query Match: 26.4% Indels: 18
 DB: 8 Gaps: 5
 US-10-780-703-2 (1-379) x ADA70350 (1-950)
 QY 52 MetProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheThrLeuArgAsp 71
 DB 52 CTCCCGGGCTTCGGTTTCCACCCAGGAGGAGGAGCTCTCGACTTCTACCTCTCCCGC 111
 QY 72 LysValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyr 91
 DB 112 GTCGTCTTAGCAAGAGCTCCACTTCAACATCATCGGCACCTCAACATCTACCGCAT 171
 QY 92 AspProTTPGluLeuProAlaMetAlaIleGlyGluLysGluTyrPheThrVal 111
 DB 172 GATCCCTGGGACCTCCAGGATGCGAAGATCGGGGAGAGGAGTGGTACTTCTTCGTG 231
 QY 112 ProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyr 131
 DB 232 CCGCGGAGCAGAGAGCGCGGAAACCGCGCGCGCGCGAACCAGCAGCAGGAGCGCGGTTC 291
 QY 132 TrpLysAlaThrGlyAlaAspArgMetIleArgSer-----GluThrSerArgProIle 149
 DB 292 TGGAGGCGCAGCGGTTCGACAGAGCGGATCCGGAGCTCCGGCGACCCGAAGCGGGTGATC 351
 QY 150 GlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSer 169
 DB 352 GGGCTCAAGAGACACGCTCGTCTTACCAGGGTCCGCCCGCGCGGCAACCAAGACGAC 411
 QY 170 TrpIleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyr----- 185
 DB 412 TGGGTGATGAAGACGATACCGCTCCCGACTACGGCGCGCGCGCGCGCGCGCGCT 471
 QY 186 GlnLysAlaGluIleSerLeuCysArgValTyrLysArg-----ProGlyVal 201
 DB 472 CCCAAGGAGGACATGGTGTCTCTGCAAGATATATACCGAAGCGCAGCAGCTGAAGAGCTG 531
 QY 202 GluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSerSer 221
 DB 532 GAGCAGAGAGCCTCGGCAATGGAAGAGATGCAG---AGAGGATCAAGTCACGAGACTAC 588
 QY 222 ThrSerSerArgLeuAlaLeu-----ArgGlnGlnGlnHis 234
 DB 589 ACGGCGCAGCAGAGCATCTCTCTCCAGCAGCGCTCCGCTCCACCGCGCAGCAGCTACTTC 648
 QY 235 SerSerSerSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnLeu 254
 DB 649 TCGTCGGACGAGCTCCACGACGAGCGGCTTCTGTATCCAGTCTCATCTCGTCGGCGG 708
 QY 255 GluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrAsnSer 274
 DB 709 CGTCCGCGCAGCAGCAGCAAGAACCGCGCGCTGGCGCACCAAGGAGGAGCAAGAGGAA 768
 QY 275 AsnSerAspValThrIleAlaLeuAlaLeu 284
 DB 769 GAAGCAGACGTCACGTCACGTCACGTCGCGTCG 798
 RESULT 15
 ID ADA45780 standard; cDNA; 1644 BP.
 AC ADA45780;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polynucleotide seqid 20520.
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;

growth rate; cell cycle pathway; disease resistance;
 galactomannan production; lignin production; plant growth regulator;
 yield; plant growth; plant development; seed oil; protein yield;
 protein content; gene; ss.
 Unidentified.
 US2004034888-A1.
 19-FEB-2004.
 28-APR-2003; 2003US-00425114.
 06-MAY-1999; 99US-00304517.
 05-NOV-2001; 2001US-00985678.
 (LIU//) LIU J.
 (ZHOU//) ZHOU Y.
 (KOVA//) KOVALIC D K.
 (SCRE//) SCREEN S E.
 (TAB//) TABASKA J E.
 (CAO//) CAO Y.
 Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 WPI; 2004-180133/17.
 New recombinant DNA construct, useful for improving plant tolerance to
 cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 pests, for conferring increased resistance to plant disease, or for
 improving yield.
 Claim 1; SEQ ID NO 20520; 15pp; English.
 The invention describes a recombinant DNA construct comprising a
 polynucleotide consisting of a sequence encoding an amino acid sequence
 available in electronic form from the US patent office at
 ftp://seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 of the invention are also useful in physical arrays of molecules and as
 plant breeding markers. The recombinant DNA construct is useful for
 improving plant tolerance to cold, heat, drought, herbicides, extreme
 osmotic conditions, pathogens or pests, for manipulating growth rate in
 plant cells by modification of the cell cycle pathway, for conferring
 increased resistance to plant disease, for producing galactomannan,
 lignin or plant growth regulators, for increasing the rate of homologous
 recombination in plants, for improving yield by modification of
 photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake,
 or by providing improved plant growth and development under at least one
 stress condition or for modifying seed oil or protein yield and/or
 content. This sequence represents a plant full length insert
 polynucleotide that can be used in the recombinant DNA construct of the
 invention.
 Alignment Scores:
 Pred. No.: 6.07e-41 Length: 1644
 Score: 535.00 Matches: 129
 Percent Similarity: 50.9% Conservativity: 73
 Best Local Similarity: 32.5% Mismatches: 130
 Query Match: 26.3% Indels: 66
 DB: 13 Gaps: 12
 US-10-780-703-2 (1-379) x ADX45780 (1-1644)
 14 SerAsnGlnValAsnAsnGluLysGlyIleGluAspAsnAspHis-ArgGlyGlyG1 33
 90 AACACATAGTGAACACACAC-----CTTGGTAATATACCATGGAGACCGCA 140
 33 nGluSerHisValGlnAsnGluAspGluAlaAspAspHisAspMetValMetPr 53
 141 AGACTCATCCAGGTTCCAGCAACACCGCAAC-----CTGCCACC 182

53 oGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTyrLeuArgAlaGlyVa 73
 183 CGGGTTTCGGTTCCACCCACAGACGAGGCTCGTGGTTACCTACCTCAGAGAGAAAGT 242
 73 lGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAspPr 93
 243 AGATTCTGTTCT 302
 93 oTTPGluLeuProAlaMetAlaAlaLeuGlyGluLysGluLysGluLysGluLysGlu 113
 303 ATGGGAATTACGACCAAGCATCGTTTGGGGCGGAGAGTGGTACTTTTTCAGCCCGAG 362
 113 gAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLy 133
 363 AGAAGGAAGTACCGAATGCGCGAGGCCCAACAGGGCTCAACTTCAGGGTATTGGAA 422
 133 sAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLy 153
 423 AGCCACGGGACAGATAAGCCTATATGTAGTGGAACTCAAAAG---GTTGGGGTGAAGA 479
 153 sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAs 173
 480 ATCTTTGGTTTCTATGGAGGAGCCACCAAGGGGTAAACCCGATTGGATCATGCA 539
 173 nGluTyrArgLeuProHis-----Hi 180
 540 TGATACCGTGTACTGAAACCAAGCCAAATAGGCCTCTCGGTGTGACTTGGGCCA 599
 180 sGluThrGluLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGl 200
 600 CAAGAAAACCTCCCTAAGGCTGGATGATGTGTGTGCGCAATCTACAAAGAGGGCAA 659
 200 yValGlu---AspHisPro-----SerValProAr 209
 660 CACAAAGGTCACATGAGAGGAGCATTCATTGATGACATGATGAGAGAGTACCTCC 719
 209 gSer-----LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSe 224
 720 TTCATCAACGTGGGCCACATGATGCGAGGTTTCACCTTTCAAGGATGTCCACGAGCTA 779
 224 rArgLeuAlaLeuArgGlnGlnHisHisSerSerSer-----SerAsnHi 240
 780 CAGTGGTGCATTGTGGAAATGACAGAAACACACCATAGAGGGGTGTATAGGCAATGG 839
 240 sSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 260
 840 TAGTGTGAATGGGATGAGCAATCTAATAATATATACAGTGTCTATCATCTCATCT 899
 260 uTyrSerGlyAspGlySer-----ThrThrThrThrThr 272
 900 GTTTGGCACCTCAAACTCCAGGCGAGAGCTTCTCTTTTGTCTTCCCTTCAAAACACATTC 959
 272 rAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIle-----TyrArgPr 290
 960 TAATTGGCGCTCCAAAAGAACACTCTCATCTCTATTTGGAAATGTGTGATGATCACAACA 1019
 290 oMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAspAs 310
 1020 CTTCAATTGGACAGATGAATGGGAATGTAGTGTGTGGAGGACAGATCAAGAGAAATAATAA 1079
 310 pAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGl 330
 1080 TAATAATAAGGTACTTCTGGCTCTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1125
 330 yAlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGl 350
 1126 -----CTTCCACAAACACCTTCATCTGCCCAATAGATGG-GTTACTTCGGA 1171
 350 yThrIleProAsnAsnAlaLeuTyrAspMetTrpAsnProIleValPro 366
 1172 CACCGTATCAATACAAAGGAGCAATTTGGTATGG-----TTAATTCCA 1214

Search completed: April 6, 2006, 11:03:49
Job time : 644 secs
